

Short-term Forecasting of Electricity Generation using Artificial Neural Networks (ANN)

DEEPAK NATESAN

Novembro de 2020

Short-term Forecasting of Electricity Generation using Artificial Neural Networks (ANN)

Deepak Natesan



Departamento de Engenharia Eletrotécnica

Mestrado em Engenharia Eletrotécnica – Sistemas Eléctricos de Energia

2020

Relatório elaborado para satisfação parcial dos requisitos da Unidade Curricular de DSEE -
Dissertação do Mestrado em Engenharia Eletrotécnica – Sistemas Elétricos de Energia

Candidato: Deepak Natesan N° 1161640, 1161640@isep.ipp.pt

Orientação científica: João Filipe Oliveira Nascimento,

Joao.nascimento@energiasimples.pt

Empresa: Energia Simples

Rua Aleixo da Mota, n.º86, R / C, 4150-044 Porto

Supervisão: Manuel Azevedo, mpa@isep.ipp.pt



Departamento de Engenharia Eletrotécnica

Mestrado em Engenharia Eletrotécnica – Sistemas Elétricos de Energia

2020

Dream, dream, dream. Dreams transform into thoughts and thoughts result in action.

Dr APJ Abdul Kalam.

Acknowledgements

I wish to express my sincere appreciation to my supervisor Professor Manuel Azevedo, for his guidance and encouragement, which enable to push myself hard when the road got tough. Without his constant inputs and words of wisdom the goal of this dissertation would have not been realized.

Importantly the opportunity to do an internship in a business environment, at a company like Energia Simples. I would also like to pay my special regards to Professor Jose Antonio Beleza Carvalho (Course Director) at ISEP for recommending me to Professor Manuel Azevedo. I wish to express my gratefulness to Engineer João Filipe Oliveira Nascimento for his mentorship in the development of this dissertation. His knowledge sharing along with his work experience.

Also, I would like to thank from my bottom of the heart to Engineer Jessica Lisandra Chaves for the advices proved monumental towards the success of this study. I wish to express my deepest gratitude to Engineer Aleksandra Krivoglazova for helping me in organize my dissertation work. I wish to thank Energia Simples team and its stakeholders for their prolonged support.

I am immensely grateful to my family and relatives. Their affection and scarifies were the burning sun in my heart. All of this would have not been possible without their support. Finally, I would like to thank my friends and flatmates for their kind words of motivation and inputs in formatting this thesis.

Resumo

As energias renováveis tornaram-se uma fonte substancial de energia em quase toda a Europa devido ao ambiente limpo e livre de poluição em comparação com os combustíveis fósseis que são usados principalmente para a geração de energia. Por estas razões também em Portugal a atenção para a utilização de energias renováveis tem vindo a crescer rapidamente. Na última década, muita investigação tem sido realizada sobre a previsão de curto prazo da produção de energias renováveis ao longo de um período de horas e dias na eletricidade ibérica. mercado (MIBEL).

Esta dissertação centra-se nas previsões para o dia seguinte e começa com uma visão teórica e panorâmica do competitivo mercado de energia em Portugal. A principal análise para focar a previsão do dia à frente é garantir o equilíbrio entre a demanda e o fornecimento de eletricidade e porque a energia necessária deve ser vendida contra o preço do mercado à vista do dia seguinte. Com base no estudo detalhado da literatura no setor de previsão para cada tecnologia, certos fatores podem ser encontrados, como localização geográfica, fonte de dados meteorológicos e fonte de dados históricos que são alavancados na precisão e para construir muitos modelos com variáveis de entrada variadas. Nos dias atuais, as ferramentas de previsão precisas são requisitos essenciais para o produtor maximizar seus lucros, admitindo perdas de lucro devido ao erro de julgamento dos movimentos de produção futuros, e para o consumidor maximizar suas utilidades.

A precisão dessas ferramentas de previsão depende principalmente do tipo e qualidade da informação que é posteriormente fornecida aos dados de entrada no modelo de previsão. Além disso, com base nas variáveis de entrada do estudo detalhado, como velocidade do vento, direção do vento, umidade relativa, temperatura, irradiação de pluviosidade e dados sazonais, foram considerados valiosos para desenvolver modelos de previsão. Muitas técnicas para prever a energia de produção para cada tecnologia são o método de regressão linear, a regressão multicamadas e as redes neurais artificiais (ANN) e a abordagem (Algoritmo de Levenberg-Marquardt (LMA))

Os resultados obtidos após a correlação (correlação do coeficiente de Pearson) da entrada é feed-forward para a rede neural e muitos modelos foram construídos. Esses resultados da rede neural foram comparados com a produção real dos resultados de predição previstos. A saída obtida é fornecida para técnicas de gerenciamento de erros, como erros de porcentagem média absoluta (MAPE). A avaliação da conformidade foi realizada por meio do Erro Percentual Médio Absoluto (MAPE) e Erro Percentual Absoluto (APE), e foi encerrando os modelos de previsão desenvolvidos para apresentarem uma precisão bastante aceitável. entretanto, para minimizar a redução de erros, os resultados são exibidos.

Além disso, muitos modelos de previsão foram executados e, posteriormente, comparados do menos com os bons modelos de previsão. Desta forma, ajuda a atingir modelos de previsão de produção de erro de baixo risco mais eficientes para cada tecnologia

Palavras-Chave

Mercado Ibérico de Eletricidade (MIBEL), produção diária, mercado de eletricidade, Técnicas de previsão de curto prazo, mercado competitivo, Coeficiente de Pearson, Algoritmo de Levenberg-Marquardt (LMA), Rede Neural Artificial (RNA).

Abstract:

Renewable energy has become a substantial source of energy nearly all through Europe since of a clean and pollution-free environment compared to fossil fuels which are primarily utilized for the era of energy. Due to these reasons moreover, within Portugal attention towards the utilization of renewable power has rapidly developing In the recent decade, an abundance of research has been performed on the short-term forecasting of renewable energy production over a period of an hour and day in the Iberian power market (MIBEL).

This dissertation focuses on day-ahead forecasting and beginning with a theoretical and overview of the competitive energy market in Portugal. The main analysis to focus on day-ahead forecasting is to ensure the balance between the demand and supply of the electricity and because of the energy required to be sold against at the day-ahead spot market price. Based on the detailed literature study in the sector of forecasting for each technology certain factors have been found such as geographical location, meteorological data source, and the historical data source which is leverage on the accuracy and to build many numbers of models with assorted input variables. Present days, Accurate forecasting tools are an essential requirement for the producer to maximize their profits, avowing profit losses over the misjudgement of future production movements, and for the consumer to maximize their utilities.

The accuracy of these forecasting tools is mainly dependent on the kind and quality of information which is later given to the input data in the forecasting model. Furthermore, based on the detail study inputs variables such as wind speed, wind direction, relative humidity, temperature, pluviosity irradiation and seasonal data have been found valuable to develop forecasting models. From plentiful of techniques to predict production energy for each technology are linear regression method, multilayer regression, and Artificial neural networks (ANN) along with approach (Levenberg-Marquardt Algorithm (LMA))

Obtained results after correlation (Pearson coefficient correlation) of the input is feed-forward to the neural network and many models have been constructed. These results from the neural network have been compared with actual production from predicted forecasting results. The obtained output is supplied to error management techniques such as Mean Absolute Percentage Errors (MAPE). The Accomplishment appraisal was performed through the Mean Absolute Percentage Error (MAPE) and Absolute Percentage Error (APE), and it was terminating the forecast models developed to show a very acceptable accuracy. however, to minimize the error reduction is accomplished which is exhibit the results.

Furthermore, many forecasting models were performed and later compared from least to good forecasting models. By this way helps to attain the most efficient, low-risk error production forecasting models for each technology have been achieved in results.

Keywords

Iberian electricity market (MIBEL), daily production, electricity market, Short-term forecasting techniques, competitive market, Pearson Coefficient, Levenberg-Marquardt Algorithm (LMA), Artificial Neural Network (ANN).

Table of Contents

Acknowledgements	i
Resumo.....	ii
Abstract:	iii
Abbreviations	x
1.Introduction.....	1
1.1 Background	2
1.2 Objective	4
1.3 Organization	4
2.STATE OF ARTS	6
2.1 Enlargement of Market Structure.....	6
2.2 REN-Redes Energeticas Nacionais	7
2.2.1 Production	8
2.2.2 Transport	8
2.2.3 Consumer	9
2.3 Wholesale Market	9
2.4 MIBEL-Iberian Electricity Market.....	9
2.4.1. Spot Market.....	10
2.5 OMIP.....	11
2.6 OMIE	11
2.6.1 Day-Ahead Market	11
2.6.2 Intraday Market.....	12
2.6.3 Market Splitting/Separating	13
2.6.4 Overall System Management (GGS).....	14
2.7 Retail Market	14
3.Essential for Forecasting	16
3.1.1 Secondary Reserve	16
3.1.2 Tertiary Reserve.....	17
3.1.3 Imbalance Maintenance	17
3.2 Negative Imbalance Production.....	18
3.3 Positive Imbalance Production	18
3.4 Imbalance Impact on Economic and System Security.....	18
3.4.1 Short-term security	19
3.4.2 Long-term security.....	19
3.5 Exist Deviation Remuneration Structure	19
3.5.1 Capacity Remuneration Mechanisms CRMs – Main characteristics	20

3.6. Renewable Energy Production	21
3.7. About Company - Energia Simples	22
3.7.1 Company Hike Towards Electricity Market	23
3.7.2 Company hike Towards Natural Gas.....	23
3.7.3 Company hike Towards Self-Consumption.....	23
4. Forecasting Model.....	25
4.1 Time Series Forecasting Model.....	25
4.1.1 Renewable Production Forecasting Model.....	26
4.1.2 Persistence Forecasting Model:	26
4.2 Naive Forecasting Methods	26
4.3 Moving Forecasting Methods.....	26
4.4 Exponential smoothing Forecasting Methods	27
4.5 Trend Projection Forecasting Methods	27
4.6 Regression Model.....	27
4.6.1 Linear Regression	27
4.6.2 Multivariate/Multiple Linear Regression	28
4.7 Artificial neural networks (ANN).....	30
5. Forecast Error Measures Accuracy	35
5.1 Mean Squared Error (MSE):.....	35
5.3 Mean Absolute Error (MAE):	36
5.4 Mean Absolute Percent Error (MAPE):.....	36
6. Study the Attributes of Energy Production.....	37
6.1 Introduction.....	37
6.2 Data Collocation and Processing	37
6.3. Correlation Analysis	40
6.3.1 Pearson Correlation Coefficient	40
6.4 Meteorological Data.....	41
6.4.1 Historical Variables based on Technology.....	42
6.4.3 Correlation of Historical Variables for All Technology	43
6.4.4 Meteorological Variables for Hydro Technology.....	44
6.4.5 Correlation of Meteorological Variables for Hydro Technology.....	45
6.4.6 Meteorological Variables for Solar Technology	46
6.4.7 Correlation of Meteorological Variables for Solar Technology.....	48
6.4.8 Meteorological Variables for Wind technology.....	50
6.4.9 Correlation of Meteorological Variables for Wind Technology	51
6.5 Section Summary	53

7. Neural Network Design and Testing	54
7.1 Forecasting Model Evolution	59
7.1.1 Forecasting Model Training Error or Overfitting	59
7.1.2 Model Parameters	60
8. Technology's and Respective Results	64
8.1 Hydro Technology Input variables	64
8.1.1 Hydro Technology Input Variables	64
8.1.2 Hydro Technology Result and Discussion	67
8.1.3 Hydro Technology Result for Hourly Model	67
8.1.4 Hydro Technology Error Obtained (Hourly Model)	69
8.1.5 Hydro Technology Results for Daily Model	71
8.1.6 Hydro Technology Error Obtained (Daily Model)	72
8.2 Solar Technology	74
8.2.1 Solar Technology Input Variables	74
8.2.2 Solar Technology Result and Discussion	78
8.2.3 Solar Technology Result for Hourly Model	78
8.2.4 Solar Technology Obtained Error for Hourly Model	80
8.2.5 Solar Technology Result for Daily Model	82
8.2.6 Solar Technology Obtained Error for Daily Model	83
8.3 Wind Technology	85
8.3.1 Wind Technology Input Variable	85
8.3.2 Wind Technology Result and Discussion	87
8.3.3 Wind Technology Result for Hourly Model	88
8.3.4 Wind Technology Obtained Error for Hourly Model	89
8.3.5 Wind Technology Result for Daily Model	91
8.3.6 Wind Technology Obtained Error for Daily Model	92
8.4 Risk Magement control for Hydro Technology	94
8.5 Section Summary.	94
9 CONCLUSIONS	95
9.1 Final Conclusions And Main Contributions	95
9.2 Future Work	97
10. References	98
ANNEX 1	101
ANNEX 2	105
ANNEX 3	110

Table of Figure

Figure 1-Scheme representation of the main objective	4
Figure 2-Evolution of the electricity sector in Portugal	6
Figure 3-Vertically Integrated Electrical System.....	7
Figure 4-Vertically integrated monopoly with competition at the production level.....	9
Figure 5- Production formation in the daily market	11
Figure 6-Intraday Market Timing.....	12
Figure 7-Market Splitting[9].....	13
Figure 8- simplified scheme two types of Market	15
Figure 9- classification of security supply	18
Figure 10-Historical development of Simple Energy	22
Figure 11-Electricity Distribution of 2018 business areas[9]	23
Figure 12-Percentage of Installed Power by Energia Simples	24
Figure 13-Quantitative forecasting classification.....	25
Figure 14-Representation of the four main parts of two biological neurons.....	31
Figure 15-three-layered feedforward neural network model[39].....	32
Figure 16-Representation of an artificial neural network.....	33
Figure 17-Schematic representation of an artificial neuronal network.....	34
Figure 18 Neural Network data manger nntool	55
Figure 19 -Example of the structure of a neural network in MATLAB's nntool tool	55
Figure 20 Creation of networks parameter	56
Figure 21 Methodology developed to choose the best model.....	58
Figure 22 Example for Reggression line error	59
Figure 23 Training Parameters	60
Figure 24 Example of a regression graphical representation, from the nntool.....	63
Figure 25 Example selecting of data variables	65
Figure 26 Comparison between the results obtained and actual production values (hourly model).....	68
Figure 27 Week from January 6 to 12, 2019, with the comparison of the Actual values of production and obtained outputs (Hourly model)	69
Figure 28 Example Histogram of the Production error obtained (Hourly model)	70
Figure 29 Example Obtained output production error and standard deviation (Hourly model).....	70
Figure 30 Comparison between the results obtained and actual production values (Daily model)	71
Figure 31 Week from February 3 to 9 (2019), with the comparison of the Actual values of production and obtained outputs (Daily model)	72
Figure 32 Example Histogram of the Production error obtained (Daily model)	73
Figure 33 Example Obtained output production relative error and standard deviation (Daily model).....	73
Figure 34 Comparison between the results obtained and actual production values (hourly model).....	79
Figure 35 Week from January 7 to 13, 2018, with the comparison of the Actual values production and obtained outputs (Hourly model)	80
Figure 36 Example Histogram of the Production error obtained (Hourly model)	81

Figure 37 Example Obtained output production relative error and standard deviation (Hourly model).....	81
Figure 38 Comparison between the results obtained and actual production values (Daily model)	82
Figure 39 Week from February 4 to 10 (2018), with the comparison of the Actual values of production and obtained outputs (Daily model)	83
Figure 40 Example Histogram of the Production error obtained (Daily model)	84
Figure 41 Example Obtained Output Production relative error and standard deviation (Daily model).....	84
Figure 42 Comparison between the results obtained and actual production values (hourly model).....	88
Figure 43 Week from January 7 to 13, 2018, with the comparison of the Actual values of production and obtained outputs (Hourly model)	89
Figure 44 Example Histogram of the Production error obtained (Hourly model)	90
Figure 45 Example Obtained output production relative error and standard deviation (Hourly model).....	90
Figure 46 Comparison between the results obtained and actual production values (Daily model)	91
Figure 47 Week from February 4 to 10 (2018), with the comparison of the Actual values of production and obtained outputs (Daily model)	92
Figure 48 Example Histogram of the Production error obtained (Daily model)	93
Figure 49 Example Obtained output production relative error and standard deviation (Daily model).....	93

Table

Table 1-Types of Market and Definition.....	3
Table 2-National Capacity remuneration mechanisms (CRMs)[26].....	20
Table 3-Total power installed and number of customers in respective years	24
Table 4-Technology and Time Interval.....	38
Table 5 Weather Data specification	38
Table 6 Precipitation data lables	39
Table 7 Precipitation photovoltaic data lables	40
Table 8 Location of each technologies.....	41
Table 9 Example how the historical production (Daily) data are regulated.....	42
Table 10. Production Correlation with Previous day hourly/daily Production by Technology. .	43
Table 11-Correlation Coefficients relative to Production (Hourly) Vs Meteorological Using Pearson´s Coefficient	45
Table 12 Correlation Coefficients relative to Production (Daily) Vs Meteorological Using Pearson´s Coefficient	46
Table 13 Correlation Coefficients relative to Solar Production (Hourly) Vs Meteorological Using Pearson´s Coefficient	47
Table 14 Correlation Coefficients relative to Solar Production Vs Meteorological Irradiation Using Pearson´s Coefficient	48
Table 15 Correlation Coefficients relative to Solar Production (Daily) Vs Meteorological Using Pearson´s Coefficient	49
Table 16 Correlation Coefficients relative to Wind Production (Hourly) Vs Meteorological Using Pearson´s Coefficient	51
Table 17 Correlation Coefficients relative to Wind Production (Daily) Vs Meteorological Using Pearson´s Coefficient	52
Table 18 Sample data for selected weeks each Technology.....	57
Table 19 Results of the parameters models created.....	61
Table 20 Test results regarding the best model based on each technology and respective five networks of training used	62
Table 21. Hydro technology hourly production forecasting models (MPH) entries variable with respective to correlation value in percentage	65
Table 22. Hydro technology Daily production forecasting models (MPD) entries variable with respective to correlation value in percentage	66
Table 23 Hydro Technology final model dropped values from	67
Table 24 Solar technology Hourly production forecasting models (MPH) entries variable with respective to correlation value in percentage	75
Table 25 Solar technology Daily production forecasting models (MPD) entries variable with respective to correlation value in percentage	77
Table 26 Solar Technology final model dropped values from	78
Table 27 Wind technology Hourly production forecasting models (MPH) entries variable with respective to correlation value in percentage	85
Table 28 Wind technology Daily production forecasting models (MPD) entries variable with respective to correlation value in percentage	86
Table 29Wind Technology final model dropped values from	87

Abbreviations

ANN	Artificial Neural Network
AvgH	Avg Humidity
Avg- $H_{b(n)_m}$	Monthly beam (direct) irradiation on a plane always normal to sun rays (kWh/m ² /mo)
Avg-Hh	Irradiation on horizontal plane (kWh/m ² /mo)
Avg- H_{i_m}	Irradiation on plane at angle (kWh/m ² /mo)
Avg- $H_{i_{opt}}$	Irradiation on optimally inclined plane (kWh/m ² /mo)
AvgI	Avg irradiation
AvgP	Avg Pluviosity
AvgT	Avg Temperature
AvgWD	Avg Wind direction
AvgWS	Avg Wind speed
EPEX	European Power Exchange Spot Market
EU	European Union
EWM	Exponentially Weighted Methods
H	Relative Humidity/ Humidity
I	Short-wave irradiation/ irradiation
LMA	Levenberg-Marquardt Algorithm
MAPE	Mean Absolute Percentage Errors
MaxH	Max Humidity
MaxI	Max. irradiation
MaxP	Max. Pluviosity
MaxT	Max. Temperature
MaxWD	Max. Wind direction
MaxWS	Max. Wind speed
MinH	Min Humidity
MinI	Min irradiation
MinP	Min Pluviosity
MinT	Min Temperature
MinWD	Min Wind direction
MinWS	Min Wind speed
NEMOs	Nominated in Each Member of State
NN	Neural Networks

OMIE	Operator-Spanish Division
OMIP	Operator-Portuguese Division
P	Rainfall Snowfall/Pluviosity
PCC	The Pearson correlation coefficient.
PCR	Production Coupling of Regions
PGIS	Photovoltaic Geographical Information System
REN	Redes Energéticas Nacionais
RM	Regression Models
SES	simplest exponential smoothing method
STFM	Short Term Forecasting Model
T	Temperature
TS	Time Series
TSO.	Transmission System Operator
WD	Wind direction
WS	Wind speed

1.Introduction

The increase in electricity demand, due to the drastic increase in population around the world which causes dioxide carbon emission affects global warming. One of the key objectives of 7th environmental action programme is to “turn the European Union into a resource-efficient, green, and competitive low-carbon economy”, also one of the horizontal priorities is to “help European Union address the international environmental and climate challenges more effectively” by 2020.[1]. Due to these action plans the renewable energy sector is growing at a blistering rate also the rapid rate of fossil fuel depletion is a vital reason. special in Portugal according to (REN) predicting by 2050 renewable energy production will fulfil electricity in the mainland of the country [1].

The data from Portuguese of renewable energy reveals that as far as power consumption in 2017 is considered, 44% of total power requirement was satisfied by renewable sources, in which the wind energy sector leads other renewable competitors by producing 11.9TWh while hydropower produces 7.3TWh[2]. In the global scale, the rise in renewables was fulfilled by unprecedented growth in China and the United States of America, accounting for around 50% rise in power production. electricity from renewable energy sources in the European Union (EU) rose by 8% and by 6% in India and Japan, respectively. Around the world, the electricity industry is lead forward towards a competitive market environment by overcoming the traditional monopolistic backdrop of the electricity industry.

The main motive of this working thesis presented here is related to study and build model short-term forecasting for the daily market production of MIBEL. This day-ahead forecasting is to ensure the balance of the demand and supply chain of the electricity. Due to demand in energy required to buy or sell against the day-ahead spot market price. Based on the detailed study in the field of forecasting for each technology certain factors have been found such as geographical location, data sources which is leverage on the accuracy and to construct network models with assorted input variable. The input variable was selected based on correlation which is explained detail. Form plentiful input data is utilized to develop a forecasting model for different technology it has been established mostly used techniques to predict production energy are linear regression, multi-layer regression and artificial neural network. The obtained result from the forecasting model is utilized to calculate error in that forecasting model. The forecasting model aspires is to prepare the customer and producer from losing money and to enhance human comfort. Based on the detail study and research a good forecast pays off economically not by generating huge money but rather by preventing from massive losses which eventually drives the organization forward.

1.1 Background

The Iberian electricity is well known by acronym MIBEL is an initiative of the government of Portugal and Spain. The main motive of the MIBEL electricity company is to structure the regional electricity market in the Iberian Peninsula. This cooperation has been profitable and outstanding not only for its contribution of electricity market on Iberian level, which also a significant step in the structure of internal energy market among the European scale throughout these years ongoing construction in which both Portugal and Spain countries have stable interconnection of electricity transmission and distribution between them and other European countries [3].

Iberian electricity region market allows the consumer to buy and sell electricity on a free competition basis from any producer or retailer within the regional market in Portugal and Spain. Later in the Iberian[4] electricity market due to some debate which was affecting market were discussed some protocols were changed to develop and join entities of Portugal and Spain signed an agreement in 2004 (Santiago de Compostela). The conclusion of this agreement well established a foundation which comes with the concept of a single energy market for the Iberian Peninsula, the MIBEL, in this energy market each sector is legally limited and separated in 4 areas such generation, transmission, distribution and electricity market. The entities connected to the areas of the transport and distribution operations under a regulated monopoly, due to this reason this model is very efficient and economical. Meanwhile, production and electricity trading were carried out in a competitive market thereby many competing agents operate.

Electricity production has been turned into a trading platform to sell and buy usually assumed that the pool trading and bilateral contracts trading. In the pool trading market producer and consumer submits the bids with respectively for selling or buying electricity on an interval, particularly on an hourly basis. Later a market operator clears the market by accepting the appropriate buying and selling bids.

According to European Directive 2003/54/EC which took place in the year 2003[5, 6] determined fully open competition in the area of electricity trading in Portugal, stipulated by law that from this day then, any consumer passes to be considered as eligible and are free to choose their supplier which in turn has a duty to ensure uninterrupted, transparent supply of electricity and at fair prices anyhow this stipulating by law was not implemented until 2007. Later period Iberian Peninsula regional market activity was reorganized and began on 24 July 2007. There was opening to competition in the area of energy production, ensuring that all agents that meet a set of requirements, according stipulated in the MIBEL, regulations which have the freedom to access trading platforms in a non-discriminatory manner.

In recent years a drastic effort has been made by the European Union (EU) (with the adoption of several directives) to construct an internal energy market which is defined as a single market that connects all markets scattered across Europe and it is expected that the implementation of this European market will bring liquidity, efficiency and social benefit.

This Iberian electricity market was first launched on 29 July 1998. It represents a regional electricity market with a phenomenal rise in renewable energy production that periodically drive forward the most expensive thermal power station outside the generation scheduling of the wholesale market reduce the cost thermal energy production which is comparatively higher than other renewable energy generation. MIBEL design market

establishes of single Iberian energy market operator-Portuguese division (OMIP) and market operator-Spanish division (OMIE).

MIBEL is a common model for calculating the marginal product of the market which is known as production coupling of regions (PCR)[7] withstanding this coupling at European level it can be said that even today the overwhelming majority of the electricity transaction in Iberian peninsula come from MIBEL, which in turn integrates four definitions.

Table 1-Types of Market and Definition

MARKET TYPE	DEFINITION
Ancillary Service Market	plays a vital role to maintain the balance between the producer and consumer
Bilateral Market	which consist of direct negotiation between the producers and traders of electricity for time horizon it basically monopoly type.
Future Markets	which aim to buy and sell electricity in the future
Spot Market	which have the aim to achieve the demand in the electricity market for a short time horizon of the day ahead (next day). This type of market was further classified into daily market and intraday market

Iberian electricity consumption majority of energy traded is satisfied by spot market obtained on contracting platform, which gives a brief picture of how the importance of daily and intraday markets for the entire value chain in the energy sector.

The intraday market is standardized in hourly basis within the competitive market electricity can be purchased a day before and the respective transaction is carried out within the intraday market. Weekdays and weekend intraday period that are considered as energy bid based on demand, which may climb up in the upcoming hours following the daily viability schedule which is fixed after the daily period. In recent few years, plentiful research was undergone in the topic of short-term electricity production forecasting model.

From the past valuable information on the forecasted hourly production that will be settled in the pool create the most beneficial material for any agent takes part in the electricity markets. A huge amount has been committed in the bidding procedure of the Iberian electricity market. special offerings were given to wind farms and production maker which enhanced profitable of the wind farm with storage capability and some offers are even given for bidding in micro-grid along with renewable generation. Therefore, most accurate production forecasts are preferred for power plants. According to production forecasts, primarily energy producers and distribution service, a huge number of customers from the demand-side based on their priority can reconsider their bidding policy by which help maximum profit.

1.2 Objective

The thesis focuses on the development of the forecasting model for different technology with various meteorological along with the historical values of electricity production while developing the models there will be a detail study and analysis for each technology. In this thesis, there are several other things achieved the entire process is shortlisted below

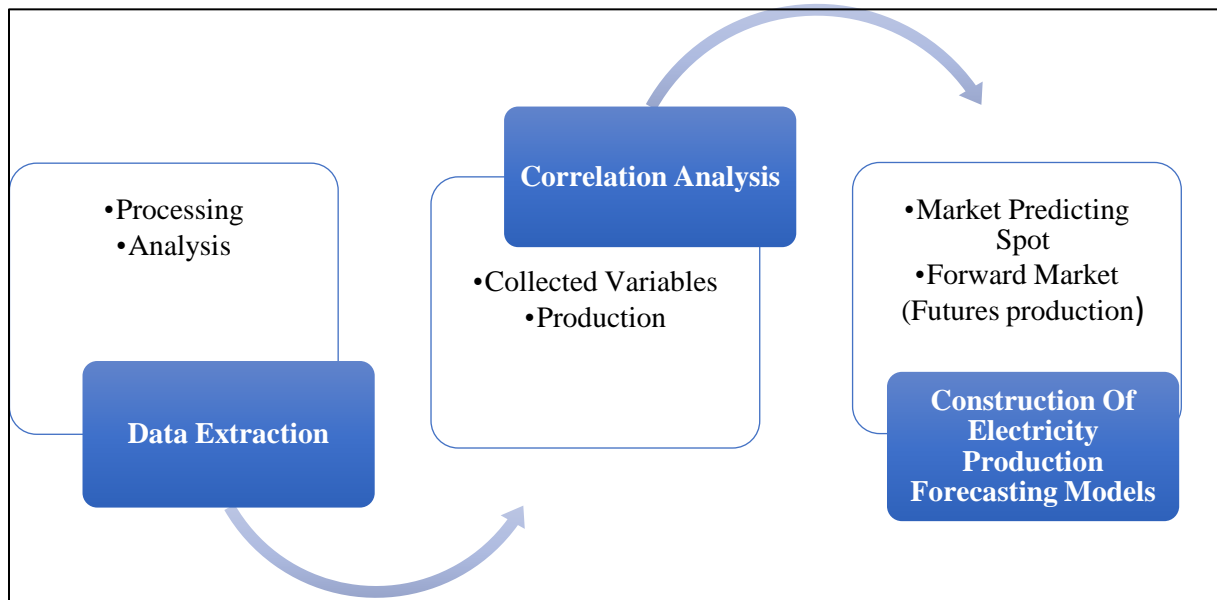


Figure 1-Scheme representation of the main objective

- I. Data extraction, processing and analysis that may influence the production of energy
- II. Correlation analysis between collected variables and production.
- III. Construction of electricity production forecasting models for the electricity market predicting spot and forward market (futures production)

1.3 Organization

This dissertation is composed of chapters, which are briefly described

Chapter 1:

Presented a contextualization of the theme addressed, the motivation for which conducted this study, the main objectives to be achieved, the timing of the work, background developed energy market, strategies, and the structural organization of the thesis.

Chapter 2:

In this section, there will be a brief discussion. study of market strategies which drives the involvement towards the market foundation to evolution competitive market in renewable production. future Portugal electricity topics like wholesale, OMIP, OMIE, and REN market strategies

Chapter 3:

In this chapter, there is a brief discussion about essential of forecasting dispute which affects forecast technique later little bit imbalance in renewable energy production in Portugal market at last about company

Chapter 4:

In this section, there is a concise analysis of renewable production forecasting model study and analysis of the various methods of existing and most used forecasting systems worldwide.

Chapter 5:

In this section, there is a concise analysis of information on the concepts and formulas for calculating deviations and their error correction

Chapter 6

In this section there is study the attributes of the energy production, where data is collected, filter and processed for each technology in a different interval of time, data collected were historical data of the production along with meteorological data of different variable based on each technology, later Pearson correlation coefficient was calculated between production and meteorological data variable.

Chapter 7

In this section there is detailed study and analysis of the Artificial Neuronal Network (ANN), by creating the network for each technology with many different variables, later forecasting model evolution with fixed parameters of nn tool

Chapter 8

In this section there is a detailed analysis of each technology along with their respective results initially stating from hydro technology there are many models were created for hourly and daily with different variables, later each network is trained and simulated output is generated later MAPE is calculated and represented in the graphical. the same thing was followed in other two technology solar and wind.

Chapter 9

In this section, there is a deep discussion of the entire dissertation conclusion

2.STATE OF ARTS

2.1 Enlargement of Market Structure

This chapter describes the evolution of the electricity sector in Portugal, going through the most important transformations that have taken place. Inception electricity market production, Portugal was massively dependent on import coal from the United Kingdom later stages it was burdensome coal import because Portugal had less heat power form coal. To overcome this problem Portugal, start, delve into main rivers for the energy production

In the year 1976, the Portuguese electric sector develops into monopoly [8-10], thus secure production, transmission and energy distribution. Subsequently with the liberalization of this sector, which helps development into a single company that could serve the market at lower production s momentous events in the evolution of the electric sector at national scheme representation is shown below. Electricidade de Portugal (EDP) came with the motivation of establishing and exploiting the public service of production, transmission, and distribution of energy in the national territory.

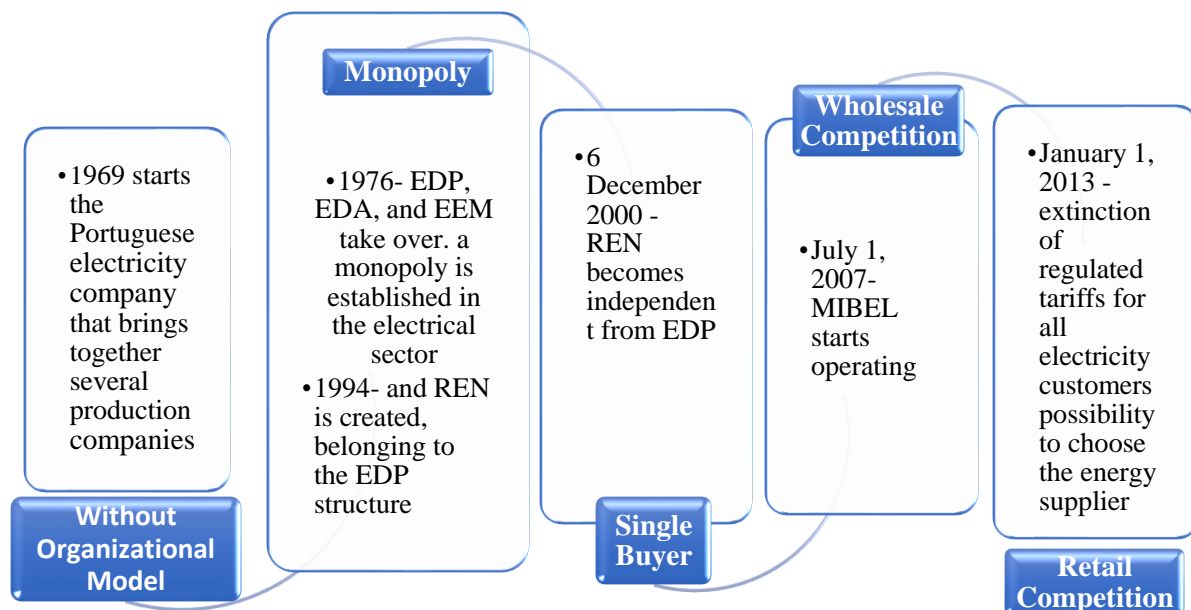


Figure 2-Evolution of the electricity sector in Portugal

In late 19th decade electricity sector in Portugal [9, 10] were driver the economy of the country by the establishment on industrialization another main objective to the transmission and distribution of electricity sector at the beginning the electricity system was composed of small networks deserted electricity lines that fed the tiny load powers. Low voltage distribution has been endorsing by municipalities by offering many go through pursue to success in

electrification throughout the national territory neglecting the efficiency issues privilege based on the contracts made between the EDP and municipalities later this system was growing in bigger as new technologies and features were implemented. This helps later to drive the electricity market and have concert connection between Spain and other countries.

Till the year 1975, the electricity sector was serene of private entities. In the same year nationalization and vertical assimilation of the sector with the creation of Electricidade de Portugal (EDP) In the same year Portugal electricity sector was nationalized [11] as well as other companies and cooperatives for the production transmission and distribution of electricity enforcing decree-law 502/76 [12]it appears that all the nationalized companies in the electricity sector are joined together in EDP(Electricidade de Portugal) in Maderia at EEM-Empresa de Eletricidade da Madeira and in the Azores, at EDA-Elétrica dos Açores EDP would then be a vertical company operating in a natural monopoly.

2.2 REN-Redes Energeticas Nacionais

REN-Redes Energeticas Nacionais [13] is culpable for maintaining the system services essential for the safe operation of the electricity system. The privilege holder of the national grid, REN is an allegation of maintaining the system service REN operates the system services market is responsible for contracting and settlement and guaranteeing a balance between the generation and consumption of electricity by means of efficient, transparent, competitive mechanism.

National Electricity Company (CNE) was established along with forerunner of Redes Energy National (REN), liable for the construction and concession of the transmission of high voltage electricity, connecting the power plant [9]. The structure of the electrical sector is shown in below Figure 3.

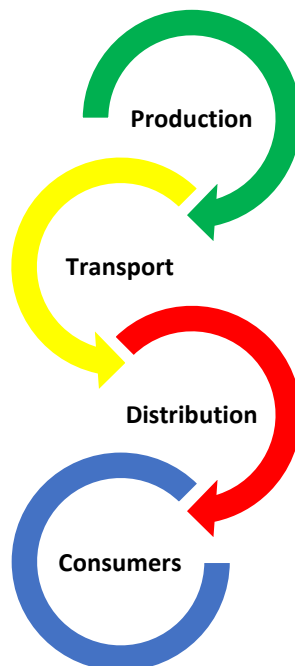


Figure 3-Vertically Integrated Electrical System

In 1970s Portugal import electricity attends to a restriction in consumption this happen due to the first oil shock that crop up in the United State of America (USA) than 1978, the PURPA (Public Utility Regulatory Policies Act) was established which reorganized electric sector where monopolistic companies have to buy energy from independent producers by this way competition in terms of electric energy production encouraging them to be more efficient.

According to the publication of decree-law no. 189/88, of 27 May, in Portugal [12], which permit any independent production to use the renewable source or co-generation, at any entity which referred to public or private thus it represented in Figure 3 of the producer in a special regime. The production of electricity under a special regime enjoyed an obligation to purchase on the part of the public network. The Portuguese state encourages an increase in the number of independent producers although limited by a set of rules which is modified over the years

Before reaching the consumer, final output production undergone various stages like production, transportation, distribution, and consumer

2.2.1 Production

It is basically generating energy from both renewable and non-renewable energy. In Portugal, electricity production is divided into two legal regimes

- ✓ **Production in Ordinary Regime (PRO)**-it is entirely related to production [9] of energy based on traditional non-renewable sources, for example, coal, natural gas, diesel, among others and huge hydro-producer centre.
- ✓ **Special Regime Production (PRE)**-related to co-generation and production energy from the used of endogenous renewable resource

2.2.2 Transport

The National Transport Network (RNT) [9] is commanded by REN in the measure of management of the electric power system (SEE), through control centres. RNT connects producer's consumption centres, ensuring a balance between energy demand and supply, thus guaranteeing the flow of electrical energy produced in power plants to distribution networks, which carry that energy to final consumers. RNT is interconnected with the Spanish network in various parts of the country, allowing exchanges of electricity with the neighbouring country, an extremely useful fact, both for the from the point of view of the safety of electrical systems, as well as from the point of supply to consumers in both countries

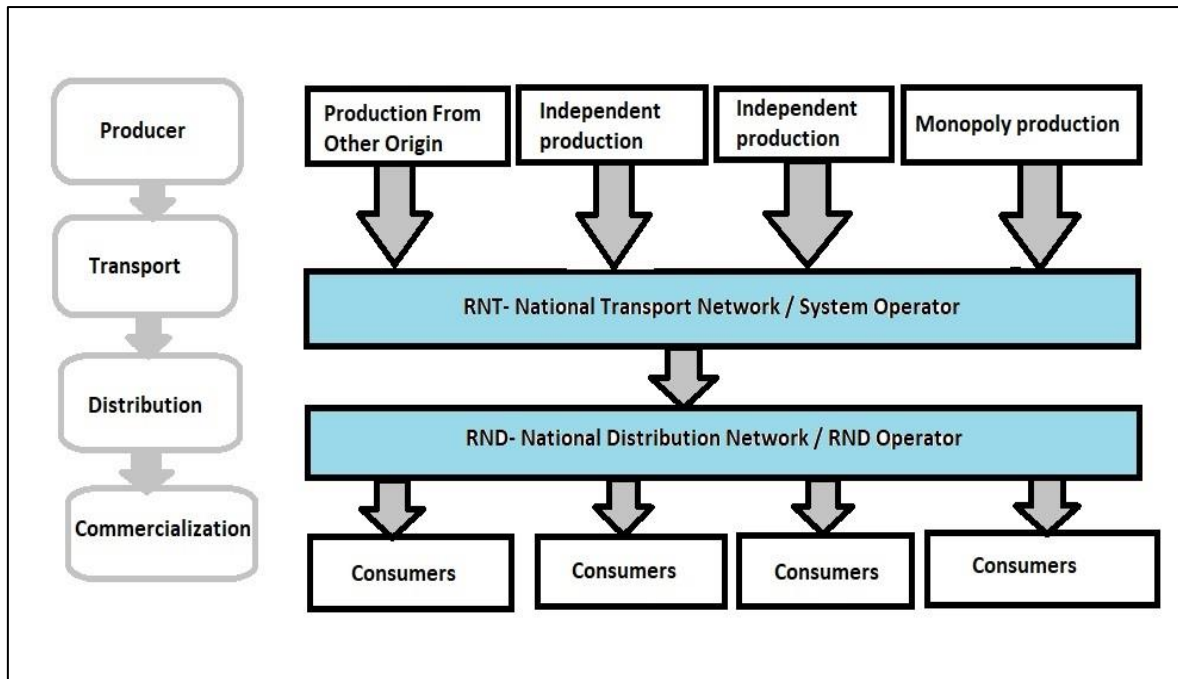


Figure 4-Vertically integrated monopoly with competition at the production level

2.2.3 Consumer

Energy trading is the last stage in the electricity sector chain. Treating currently a free activity is, however, subject to the granting of a license. In Portugal, there are more than 20 electricity suppliers that operate within the scope of the energy market. To protect consumers [9], the existence of last resort (CUR), whose objective is to guarantee the supply of electricity to consumers, namely those who do not intend to join the liberalized regime (in this case, the tariffs charged are defined annually by ERSE).

2.3 Wholesale Market

The wholesale market[9] attribute to purchase and sell the energy between the generators and resellers reseller include electricity utility companies, which is responsible for provides energy to the electricity markets.

In the past decade, the organization of the wholesale market in Spain and Portugal has forwarded into a base level of horizontal concentration even though the momentous difference between the structures of the two markets. By the regular observation of market shares and concentration indices of the generation supply both in terms of energy produced and installed capacity. In both countries, this trend is primarily related to the fact that there are new agents covered under special regime generation to a certain extent

2.4 MIBEL-Iberian Electricity Market

The government's main objective is to promote the integration in both the mainland of Portugal and Spain electricity system. The repercussion made significant improvement not only towards the implementation of the electricity market at the Iberian base but also at the European base as an important step towards establishing an internal energy market

MIBEL market [14] was developed in the year 1998 there were undergone various government-contracted during the different time period which is shortlisted below

- I. Establishing of Iberian Electricity market by the government in the mainland of Portugal and Spain this protocol was a sign in November 2001.
- II. Agreement in Santiago de Compostela between the two countries signed in October 2004
- III. 22nd Luso-Spanish summit of Badajoz which took place in November 2006 and in January 2008.
- IV. The Braga Agreement which revises the Santiago agreement were signed

The system framework would bring prosperity for the consumer of both countries which providing access to all attentive parties in terms of equality, transparency, and impartiality all this adjustment was made and finally on 1 July 2007 MIBEL was successively launched

The electricity market can be classified into two categories the spot market where energy is traded for prompt physical distribution whereas in the future market the distribution is done at later period and commonly do not relate to physical distribution. The future market is commonly used for risk fudge. Regularly European wholesale market is classified into two days ahead and intraday session which is briefly discussed upcoming section

2.4.1. Spot Market

Unlike most other merchandise or financial market, the electricity spot market[15] is routinely a day-ahead market that does not allow for stable trading. This is a consequence of system operators requiring prior notice to certify the schedule is achievable and drop within the transmission constraints. In the day-ahead market, agents submit they are and offers for the transmission of electricity production during each hour (shorter period) of the next day before market concluding time. Thus, when dealing with the modelling and intraday forecast of electricity production, its critical to lock-back that, in most markets production for all the contracts of the next day are determined at the same time using the data available.

The palpable role of a structured market for electricity production is to have a power exchange or power pool within the market structure this is done to balance supply and demand and determine the market-clearing production price. (MCP). Routinely the MCP is entrenched is the one-day valuation of the intersection between the supply curve and the demand curve electricity production is increasing or decreasing until the total demand-supply is met

In some accusation, if there is no production transmission bottleneck, then MCP is the only production price for the entire system nevertheless when there is bottleneck locational marginal production price (LM). Nodal production is the sum of generation marginal and transmission congestion cost it can be diverse for diverse nodes, even within a local area. They are the quintessential reference because the electricity production value is based on where it is generated and transmitted. however, they routinely drive to a higher transaction and greater complexity of the pricing mechanism. These productions can be the same within the area and varying from one area zone to another zone.

2.5 OMIP

The Iberian Electricity Market (MIBEL) is an elbow excitement of the Portuguese and Spanish government with a glimpse to the development of a provincial electricity market. With MIBEL subsequent into force, every consumer in the Iberian Peninsula was able to obtain electricity, under free competition, from any producer or seller operating in both Portugal and Spain. A term contract market (OMIP) [9], in which commitments have established the future of production and purchase of electricity. This market can physical settlement (delivery of energy) or financial settlement (compensation of monetary values underlying trading).

2.6 OMIE

OMIE is represented as electricity market operator of Nominated in Each Member of State (NEMOs) for maintaining the Iberian Peninsula's day-ahead and intraday electricity market OMIE [16] also participates in connecting with wholesale electricity market in the European Union along with all members of NEMOs. Europe has established a regulatory background for the European electricity sector until 2030 based on the marginal cross-border energy market. Under this regulation, OMIE manage the day-ahead and intraday wholesale electricity market for Spain and Portugal

2.6.1 Day-Ahead Market

In this market, an aggregator can be able to buy or sell the energy for the next 24hours in the closed auction. The offer made by the aggregator to buy or sell should match to market production otherwise. purchase or selling the energy will be declined in the closed auction. The day-ahead has been drastically dependent on demand and supply. This day-ahead market was implemented since the year 1998 January 1st in Spain and in the year 2007 July 1st in Portugal. In these first 18 months that the market splitting mechanism has operated trading volume has outpaced about 374.4TWh for Iberian system fact [5, 17]. Later, the average Spain-Portugal hourly production spread has been deliberately decreasing.

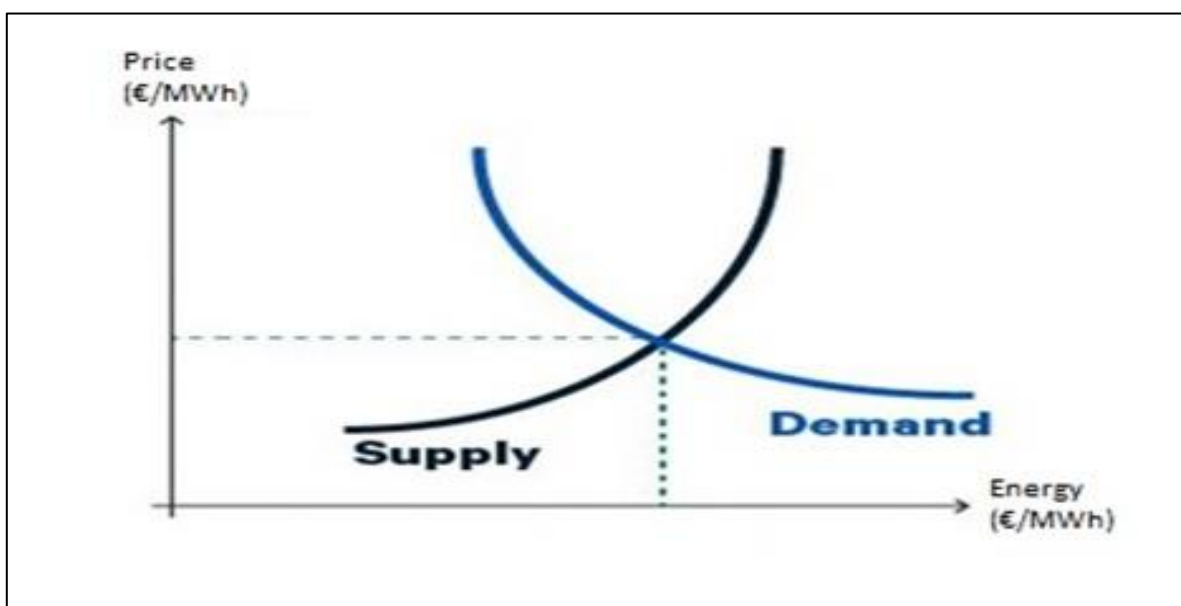


Figure 5- Production formation in the daily market

2.6.2 Intraday Market

Intraday market [17] trading is continuous purchase and selling of energy which takes place on the same day as the power delivery. The largest intraday power exchanges in European are the EPEX spot (European Power Exchange Spot Market) [18] This market is established works together with the day-ahead to necessary balancing the demand and supply of Iberian market. This market gives the operation and optimization multilevel flexibility of the portfolio of agents in successive series of time horizons and offer is similar guarantees in terms of transparency and monitoring as the day-ahead market.

In the intraday, electricity is traded to modify the quantities traded in the daily market, comprising six daily trading sessions. Each of these intraday market sessions configures the production for the hours liable to each session, as shown in Figure 6

A small brief discussion of the intraday session regarding Figure 6

- ✓ The first intraday session forms a production price for the 24 hours of the day following the negotiation.
- ✓ The second intraday session is production price for the last 4 hours of the trading and for the 24 hours of the day following the trading.
- ✓ The third intraday session forms a production price for the 24 hours of the day following the negotiation.
- ✓ The fourth intraday session forms a production price for the 20 hours between the hour 5 and hour 24 of the day following the negotiation.
- ✓ The fifth intraday session forms a production price for the 17 hours between the hour 8 and hour 24 of the day following the negotiation.
- ✓ The sixth intraday session forms a production price for the 12 hours between the hour 13 and hour 24 of the day following the negotiation.

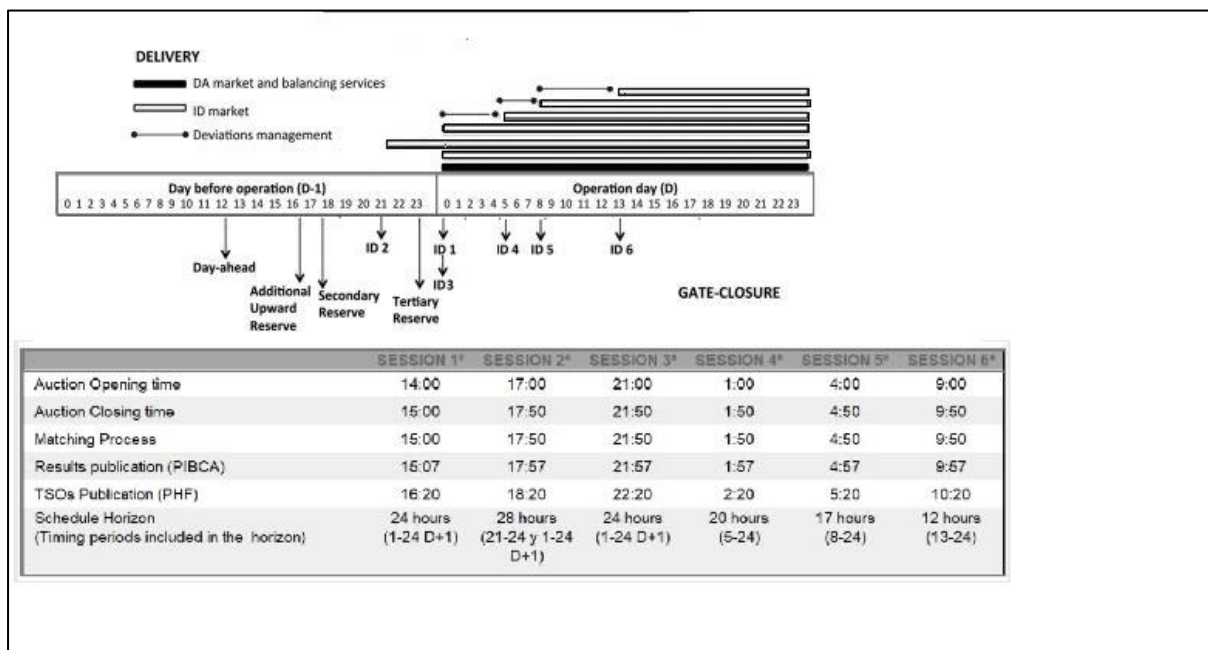


Figure 6-Intraday Market Timing

2.6.3 Market Splitting/Separating

The integration of markets between areas/countries there is a complication that stands out transmission capacity available to transport energy from one destination point to other destination This mechanism method was introduced since 2007 to connect a strong bond in the market between Portugal and Spain. This interconnection consists of a market separating [5, 9] the mechanism in the daily horizon that keeps exhausted all available capacity safely.

The fact that the spot market does not have the transport production which means that the global dispatch of the system may not be technically feasible, and trade must be changed. Thus, with interconnection lines, the surplus energy on one side and the energy deficit on the other side must be balanced by making slight adjustment which creates new dispatch is made for various areas despite the various production. If in case, there is free capacity in the interconnection the production will be the same for the different areas.

This mechanism established a marginal production method in which every individual area act as a bus. In case the power supply is not carried out at minimum cost then energy will have to be supplied from expensive generators

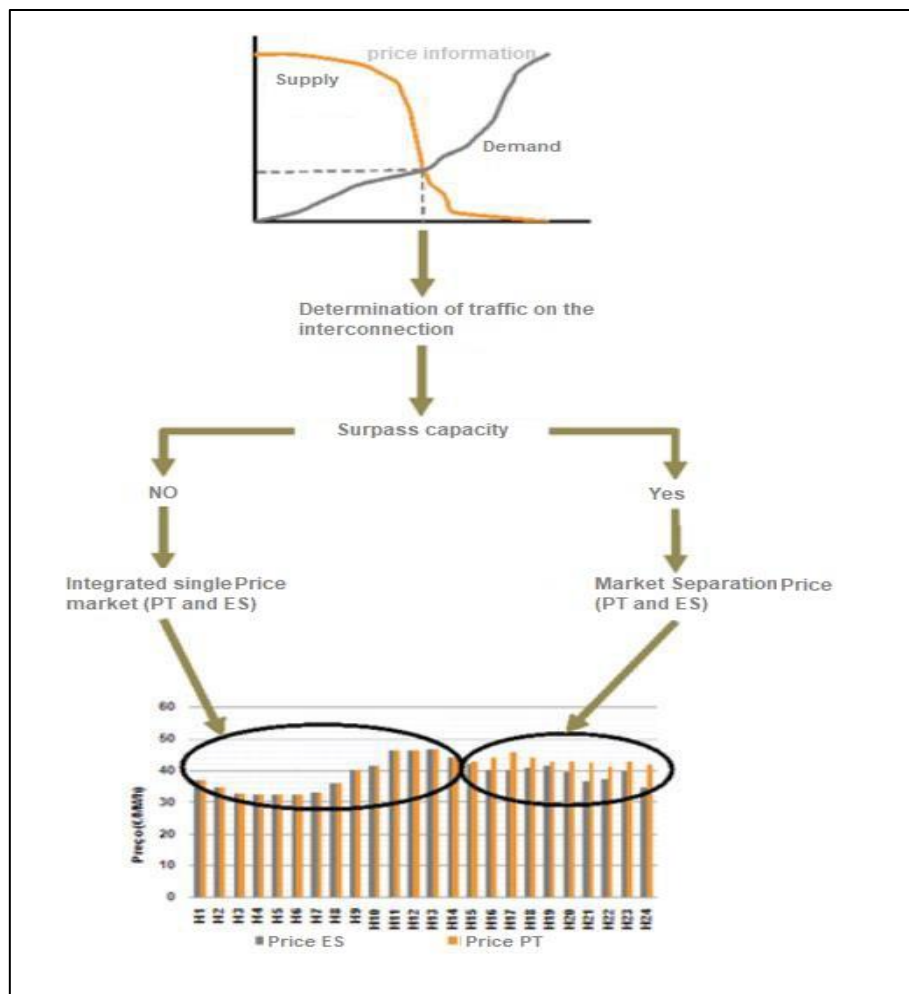


Figure 7-Market Splitting[9]

2.6.4 Overall System Management (GGS)

Overall System Management (GGS) [19] helps to keep demand and production balance according to GGS procedures manual they are classified into three reserves they are shortlisted below

I. Primary Reserve

The response of the turbine speed regulator due to a shift of the electrical system regularity with the aim to modify the generator power. The conclusion power variation should take within 15-30 second for frequency shifts of 100-200mHz

II. Secondary Reserve

This reserve is used to maintain the shifts in the interconnection within Spain to collude in sustain the joint frequency this reserve should not take more than 30 seconds and response should also be completed with the time period eventually completed action of the regulation should not exceed 5 minutes.

III. Regulating Reserve

The maximum power difference between the system variation can be used within the current operation programming period. In some circumstance where monitored previous production in the last program is varying from GGS demand forecast, then mobilization or demobilization of the production/consumption takes place in order to balance production with the referred to the demand

2.7 Retail Market

The retail market[9] has a greater level of concentration, related in part to the integration between distribution and marketing of existing operators, which has historically led to customers remaining with providers within the same business group. Likewise, the development of marketing and the emergence of new competitors have been limited in recent years due to the problem of the tariff deficit, causing consumers' participation in the free market to be delayed. The retail market is classified into two main forms of contracting the supply of electricity to end consumer

- ✓ Contracting in a regulated market by applying full tariffs regulated
- ✓ Contracting in the liberalized market with the condition for trading energy to be defined and agreed between the parties and the network access component to be applied through regulated production.

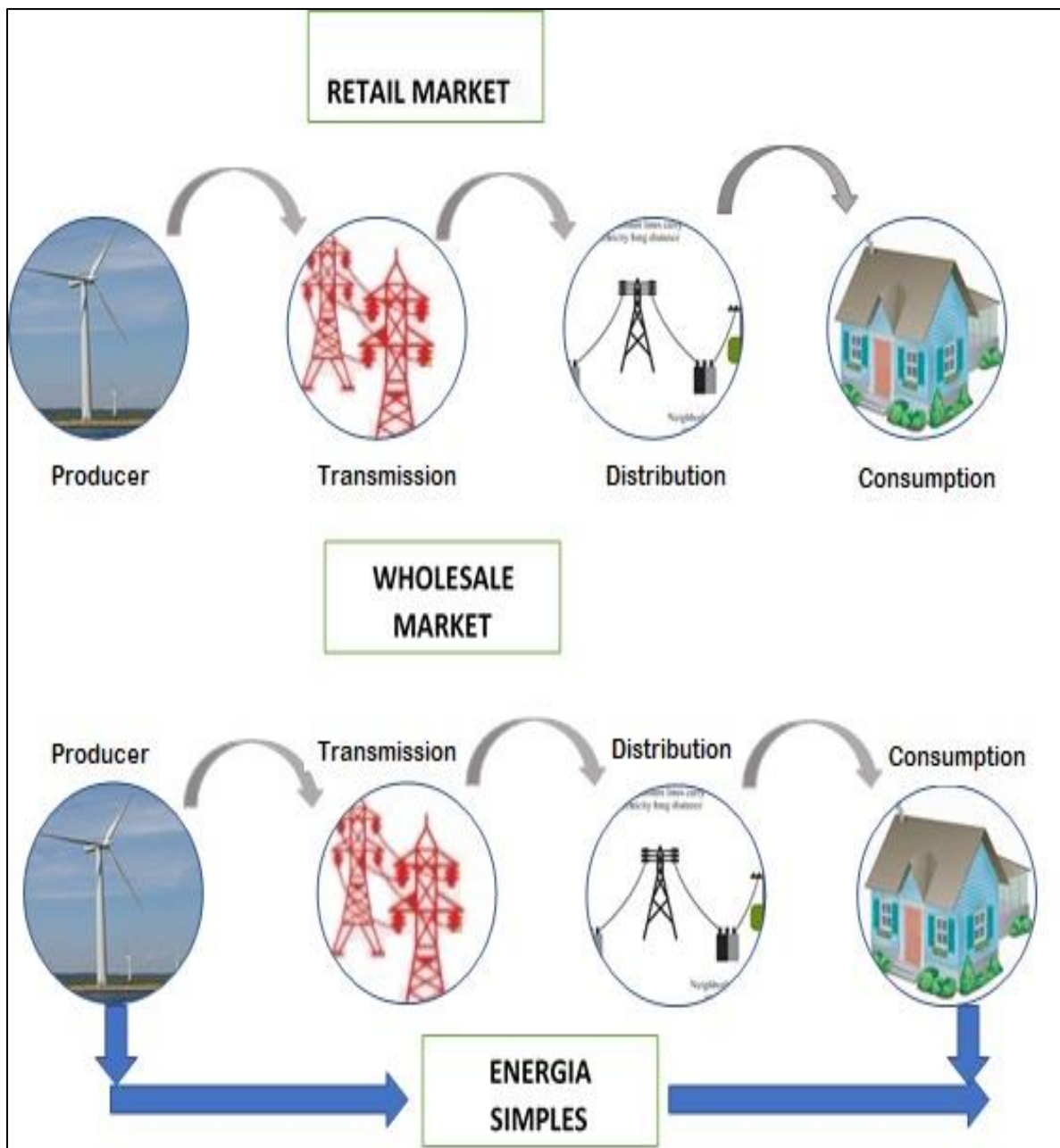


Figure 8- simplified scheme two types of Market

3. Essential for Forecasting

While the forecast is never been flawless but still it's essential for the forecast to prepare for actual demand in order to maintain an optimized inventory and effective supply chain and accurate demand. Since the early 90s the process of the deregulation and introduction of the competitive market has reshaped the landscape of this traditionally monopolistic and government control power sector in many countries worldwide. Electricity is now traded under market rule using spot and derivative contracts. However, electricity is a very special quality, is economically unstable, power system stability requires a constant balance between the production and consumption at the same time. Electricity demand on the weather, temperature, wind speed, precipitation, Cetera, and intensity and everyday activity, on the one hand, these unique and specific characteristics lead to a production that is not observed at the other market exhibiting seasonality, the daily, weekly, and annual levels, and abrupt short-lived unanticipated production spikes. On the other, they have encouraged researchers to intensify the efforts in the development of better forecasting techniques. At the corporate level, electricity production forecasts have become a fundamental input in the energy company's decision-making mechanism.

Electric utilities are the most vulnerable since they usually not able to pass production on to retail consumers. The production of over or under contracting and then selling or buying power in the bouncing market are typically so high, they can lead to huge central losses. Forecast for few hours to few days which is basically short-term forecasting [15] ahead of the kind of particular interest to power portfolio managers, a generator utility company or large industrial consumers, able to forecast the volatile wholesale production with reasonable accuracy can adjust its bidding strategy and its own production and it helps to schedule and reduce the risk and maximize the profits.

3.1 Imbalance Market

During the interval between last closing session of an intraday market session and opening of the next day along with deviation between of scheduled and measured energy are forward through ancillary services is based on the market producers, such as secondary reserve, tertiary reserve and imbalance [20] management process.

3.1.1 Secondary Reserve

The secondary reserve [20] is an optional ancillary service. The motive of this procedure is two-fold: to compensate for the power deviations, sustain the power system regular reference value. This ancillary service consists of supervision zones, serene of many generators put together in a zone and referred to as energy scheduling units which are responsible for achieving the commitment of the Transmission System Operator (TSO). The payment was done by two concepts of market procedures: stock availability i.e. secondary reserve and utilization of energy. The market deal with generators bid both in rising and fall of secondary reserve power and their corresponding production s . These bids or offers are allocated using minimum production considering system constraints, to obtain a rise or fall of production in the secondary reserve.

3.1.2 Tertiary Reserve

This ancillary service is maintained according to market procedures. Even though it is mandatory only previously authorized generators can engage in the tertiary reserve [20] markets. This service objective is to restore the secondary reserve band. The marginal up/down productions of the entire tertiary reserve market are gathered from the allocation of both up and down bids and their respective productions using minimum costs producers moreover power system constraints could be altered the initial market schedule.

3.1.3 Imbalance Maintenance

The imbalance maintenance process[20] is a choice of service main objective is to determine the difference between generation power and energy demand that may occur at last closing session of one intraday market until the first period of the horizon of the next intraday market session conventional power producers, pump storage and manageable renewable energies can engage in markets. The TSO can offer the bids is drastically depending on the cost for forecasting imbalance between generation and demand. During this process, the energy scheduling units that can participate in the surplus market bid their rise and drop in energy program once the bids have been allocated the production for all scheduled modifications is the marginal product of the short /long imbalance.

In communally the balance responsible party [21] is the entity culpable for the imbalance generated by deviations from predicted schedules. Within the EU each country electricity market situation with a considerable power imbalance

Great Britain- In GB [21] BETTA is responsible for electricity trading. This trading is done through bilateral contracts between producers' distribution and customer across a series of markets operating Both energy producers and suppliers is essential to provide forecasts of their planned demand and producer until respectively by gate closure time TSO responsible for buys balancing service from producer and load customers to balance demand and supply in case of any deviation from schedules are traded at balancing market production.

The Netherland-In the Dutch liberalized energy [21] market balancing is a liability on Programme Responsible Parties (PRP_s). Every generator and load in the Netherlands are assigned to a PRP which is responsible for the TSO to maintain the scheduled quarter-hourly energy exchange with the Dutch system of all generators and demand in their portfolio Deviation from the schedule are penalized by the TSO.

Germany- In the Germany energy market, four TSOs take care of the balancing in their control areas. The unique German imbalance production [22] (reBAP) is the settlement production for all balancing Responsible Parties (BRP) in Germany. Next to this purpose, the reBAP has a fundamental function for the German electricity market. It is up the reBAP to set correct subsidies forwarded to BRP to close a position in the market before TSO need to activate balancing energy to balance open positions

Spain-In the Spanish electricity market[23] there is classification know as Red Electrical[24] which is responsible for the system operator consists of ensuring balance in the Spanish Electricity system. All the renewable energy producers are responsible for the

production of any schedule deviations and for paying the production of the balancing energy demand.

3.2 Negative Imbalance Production

Negative imbalance production [25] is a remedy signal representation of the systems famine of downward resilience but also depend on market distortion caused by renewable support mechanism. In the calamity situation, these hold active market participation for example in Germany, with its feed-in tariff coupled with a curtailment policy. In the best situation, these result in negative bids under their marginal production to recover lost subsidies, as in Belgium.

3.3 Positive Imbalance Production

Positive imbalance production [25] is a remedy signal representation of the systems famine of upward resilience, but also depend on market distortion caused by renewable support mechanism.

3.4 Imbalance Impact on Economic and System Security

System Security [26] of electricity supply is the ability of the electrical power system to provide electricity to end-users with a specified level of continuity and quality in a sustainable manner, relating to the existing standards and contractual agreements at the point of delivery. Security of electricity supply derives from the interaction of several factors as

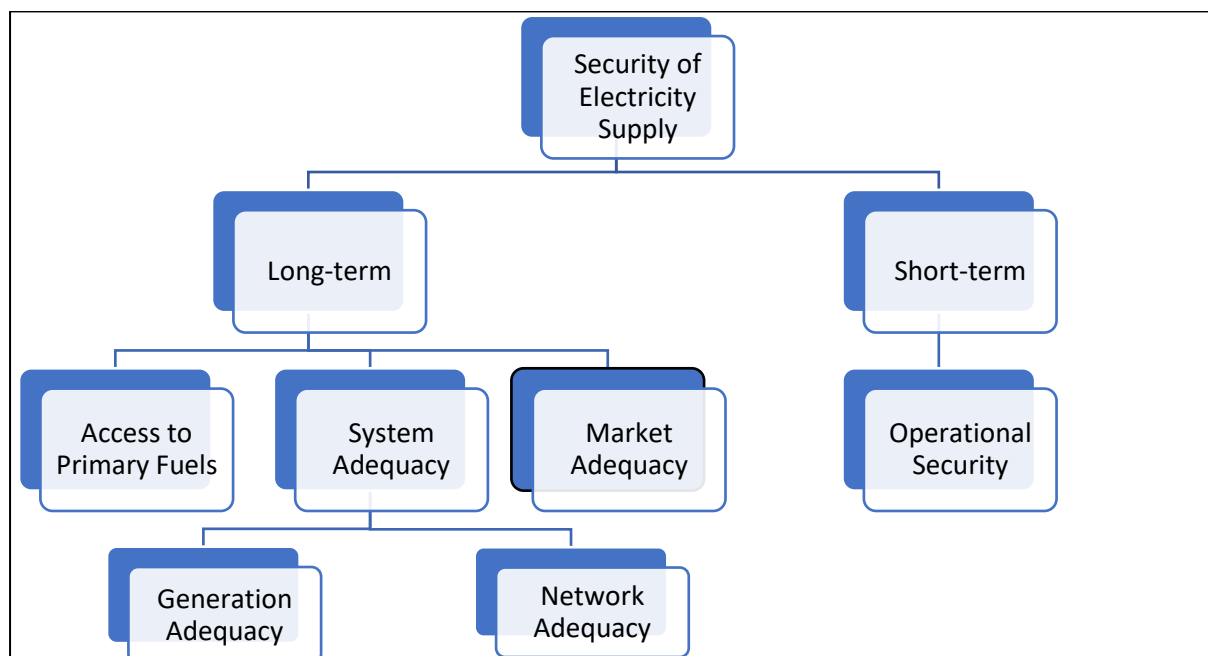


Figure 9- classification of security supply

3.4.1 Short-term security

Short-term security [26] of electricity supply refers to the operational reliability of the system as a whole and its assets, including the ability to overcome short-term failures of individual components (generation facilities, power lines, grid equipment, etc.) of the system. Operational reliability also refers to specific cases where system security is threatened by grid-related physical congestions border trade at interconnections are often. A significant share of the physical capacity is kept as a precaution or safety margin for possible outages and loop flows (Transmission Reliability Margin)

3.4.2 Long-term security

Long-term security [26] of electricity supply is the simultaneous adequacy of access to primary fuels, generation, networks and markets. Access to primary fuels means that electricity producers are allowed to choose System adequacy is the ability of the electricity system to convert primary fuels into electricity and transmit that electricity to end-users in a sustainable manner the electricity system needs to be able to provide sufficient electricity to meet demand at all times and in all parts of the system

In most cases, long-term security of supply depends on market design and business decisions of investors Regarding the issue of market design, Insight research on power plant dispatch at EU level shows that the integration of renewables may lead to lower remuneration of conventional power plants¹⁶, preventing the formation of an electricity rent, i.e. a contribution margin component reflecting a production signal to support long-term investments in new facilities or refurbishment.

Therefore, in the context of generation adequacy, Capacity Remuneration Mechanisms (CRMs) are discussed as a supplement to Energy Only Markets (EOMs). An Energy Only Market (EOM) procures the remuneration of an energy component (kWh) on the wholesale electricity market under different time scales. Both CRMs and EOMs constitute different forms of electricity market designs [26]

3.5 Exist Deviation Remuneration Structure

Europe has appreciated a secure electricity supply for decades associated with very high grid security standards. In the context of liberalization and the creation of a European electricity market, security conditions underlying the supply of electricity need to be reconsidered. and the growth in the share of renewable energy sources in electricity generation that is projected.

According to the European Commission's 2030 energy package, European legislation on Security of Supply (Directive 2005/893 or Security of Supply Directive) sets forth a general definition of security of electricity supply - "the ability of an electrical system to supply final customers with electricity" - and provides the foundations of security of supply at Member States' level based on a long term projection of electricity generation capacity and demand estimates (generation sufficiency plans)[26]

3.5.1 Capacity Remuneration Mechanisms CRMs – Main characteristics

Capacity remuneration mechanisms[26] are generally introduced to overcome one or more shortcomings of energy-only markets. As the name suggests, CRMs foster investments into new generation capacity. EU Commission guidance on public intervention provides the member states with the possibility to implement a CRM if there is evidence that the new scheme will support interconnections and storage

European legislation, therefore, recommends demand response or storage units to be activated as sources of flexibility to complement intermittent generation. CRMs are clustered as volume-based schemes (capacity auction, subscription, obligation, reliability option, capacity reserves) and production-based mechanisms (capacity payment). Capacity reserves allow covering fixed/variable costs independently from the EOM.

Table 2-National Capacity remuneration mechanisms (CRMs)[26]

Country	Impact on Capacity remuneration mechanisms
Italy	<ul style="list-style-type: none"> • Targeted capacity payments for dispatchable generators • Critical periods there is additional generation capacity is expected to be required which is used in critical periods • Case of producer fails to produce the capacity NRA impose a fine amount is collected by capacity mechanism is borne by end consumers
France	<ul style="list-style-type: none"> • Decentralized capacity under this scheme French market hold certain guarantees capacity • This holding capacity is necessary calculated every year
Portugal	<ul style="list-style-type: none"> • Targeted capacity Scheme • Incentivize scheme was introduced to promote new hydro generation and repowering of exiting pump storage • Eligible power plants can only benefit from incentivizing scheme with minimum cdf • The cdf also determines the annual incentive amount to be awarded
Poland	<ul style="list-style-type: none"> • Targeted capacity mechanism • Demand response-specific capacity mechanism remunerates demand only if actual curtailments are carried out by the TSO

	<ul style="list-style-type: none"> • Moreover, during actual scarcity periods, the TSO can curtail demand administratively (without remuneration) rather than through the mechanism.
Germany	<ul style="list-style-type: none"> • activated under specific circumstances i.e. scarcity circumstances if the day-ahead market cannot be cleared. • Additional costs are accounted as grid costs for end-users. Strategic reserves have been used in other countries like Sweden
The Netherlands	<ul style="list-style-type: none"> • The Dutch government believes that capacity mechanisms are a “second-best” option. • The extension of interconnection capacity bearing in mind the highly interconnected level of the Dutch electricity market

They all include the oversight of TSOs and/or NRAs in setting either volume (capacity reserves) or the objective (adequacy, Loss of load expectation), envisaged under a market instrument. security of supply (generation adequacy) and is “technologically neutral” i.e. open to alternative options, such as demand response[26]

3.6. Renewable Energy Production

Renewable energy[9, 27] is an energy that comes from natural sources with the capacity for renewal. renewable energy technologies use resources straight from the environment to generate power. These energy sources include sunshine, wind, tides, and biomass, to name some of the more popular options. Renewable resources won’t run out, which cannot be said for many types of fossil fuels as we use fossil fuel resources, they will be increasingly difficult to obtain, likely driving up both the cost and environmental impact of extraction. renewable energy requires low maintenance compared to all other fossil fuel

The creation of these renewal sources is essentially linked to social and environmental factors. With the development of such sources, increasing global energy security; reduction of carbon emissions and the creation of new jobs. However, there are some barriers to its implementation, such as:

- I. Costs: Renewable energies need an initial investment higher than the conventional energies, that is, renewable energies generate less capacity per euro invested than traditional energies.
- II. Legal and Regulatory: Lack of legal framework for producers and the existence of restrictions on construction in terms of aesthetics, noise, safety, which can affect agricultural or environmental zones.
- III. Intermittency: Though renewable energy resources are available around the world, many of these resources are not available 24/7, year-round. Some days may be windier

than others, the sun does not shine at night, and droughts may occur for periods of time. There can be unpredictable weather events that disrupt these technologies. Fossil fuels are not intermittent and can be turned on or off at any given time.

- IV. Storage capabilities: Because of the intermittency of some renewable energy sources, there is a high need for energy storage. While there are storage technologies available today, they can be expensive, especially for large-scale renewable energy plants. It's worth noting that energy storage capacity is growing as the technology progresses, and batteries are becoming more affordable as time goes on.
- V. "Market Performance": Technological uncertainty and greater risk, and the lack of technical or commercial experience or information.

However, despite these barriers, in 2016 renewable energy represented 57% of the energy produced in Portugal, working only with solar, hydro and wind energy for 107 consecutive hours. Although some of the national energy produced was exported, Portugal consumed, in that year, 28.5% of its energy from renewable sources, being the third EU country that produced the most renewable energy.[9]

3.7. About Company - Energia Simples

Energia Simples is a Portuguese trader of electricity, natural gas, and integrated photovoltaic self-consumption solutions the main tagline of the company is "Our Dream Go Green Energy". Headquartered in Porto, Energia Simples[9] was founded in November 2014 by a group of professionals in the areas of energy, engineering and technology, but with concerns common, namely: energy efficiency, environmental sustainability and ethical business. The following Figure 10 shows the diagram of the most important events in the company's evolution.



Figure 10-Historical development of Simple Energy

Its main commercial activity, represented in Figure 11, resides in the commercialization of electricity. In precise, has extensive knowledge in the area of selling and buying of electricity in the Iberian market, in the development of indexed tariffs, in the dimensioning of self-consumption systems and also facilitator services in the energy market with the aggregation of distributed production plants and their marketing on the retail market. Through the policy of introducing innovations to the market in a sustained and continued, it was possible to develop innovative business models, ensuring a highly competitive positioning with customers most concerned with the energy consumption and environmental sustainability[9].

Energia Simples Business Areas

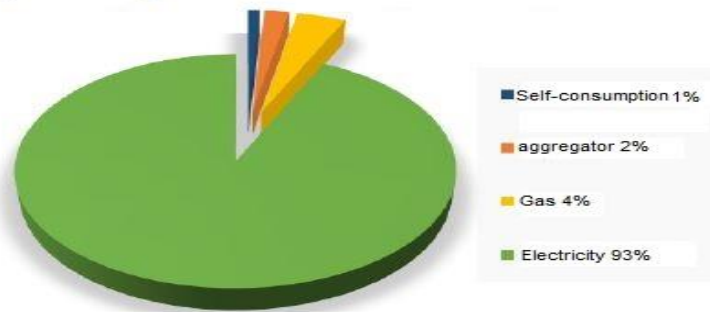


Figure 11-Electricity Distribution of 2018 business areas[9]

3.7.1 Company Hike Towards Electricity Market

The area of commercialization of electricity continues to be further momentous for the enterprise Simple Energy, having seen tenacious growth in 2017. “In 2017, Energia Simples carry out a turnover of 101.4 million euros, which allowed us to occupy the 7th position among the largest companies in the market liberalized in the consumption quota. This result is representative of the growth fortify of Energia Simples, with a total of 14,929 to date customers, 769 GWh of energy sold, 30 GWh of gas sold and a hike of approximately 148%. ”

In 2018, despite the decrease in the number of customers, figures presented at the end of the 11,925 active customers with an annual consumption of 795.45 GWh, of which drastically up towards than the previous year[9].

3.7.2 Company hike Towards Natural Gas

The commercialization of natural gas started in 2017, three years after the groundwork of the company, having in the first year of activity a total of 30 GWh of energy marketed, representing a portfolio of around 850 customers. Currently, there are 895 gas customers, with an increase of 45 customers compared to the year Figure 11 Distribution of business areas 2018.consumption and an increase of 118.78 GWh in consumption, making a total annual consumption of 148.78 GWh.[9]

3.7.3 Company hike Towards Self-Consumption

Arduous of Energia Simples, PH Solar was created, starting its enterprise in the area of energy for more than two decades, through the promotion of production units decentralized, energy monitoring and energy efficiency. The imperative evolutionary path has been verified at the business level, currently, and already with the Simple Energy, there is a customer portfolio, in an evolutionary system increasing, in the Production Units for Self-consumption (UPAC), allowing customers produce your own energy locally according to Table 3 shown below[9].

Table 3-Total power installed and number of customers in respective years

YEAR	NUMBER OF CLIENTS N⁰	TOTAL INSTALLED POWER (kW)
2016	20	344.36
2017	22	423.67
2018	31	1350.3

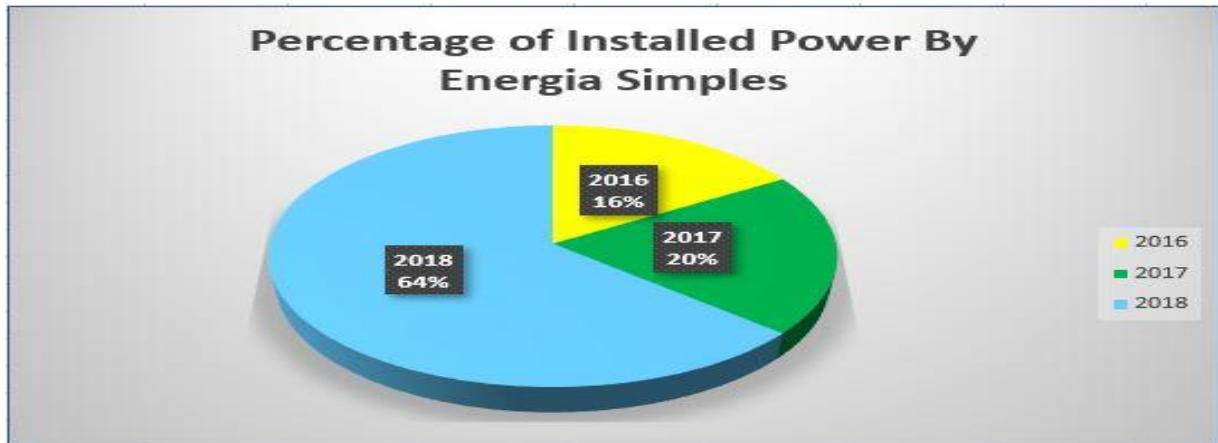


Figure 12-Percentage of Installed Power by Energia Simples

In 2018, as can be seen in Figure 12, there is an inflation of approximately 30 new customers, thus totaling 1350.3 kW of installed power.[9]

4. Forecasting Model

Different forecasting methods and in a couple of techniques that we will drive a little bit deeper into those forecasting methods, and specification. To create those models and how to determine which model is right for us or if we need a more robust model as we been looking at forecasting methods one of the first things that most research or people think of is they think about the quantitative type of forecasts which is taking the historical data of the past or time series or correlation data with creating these forecasts out into the predicted future. But there are actually quite a few instances where we see a lot of qualitative forecasting approaches and these would be opinions from expert's decision-makers or customers about what should expect to see in a specific situation and though thesis cover the quantitative of energy producer point of view it is necessary that we should other points of approach towards energy forecasting.

Quantitative forecasting methods can be broken down into four

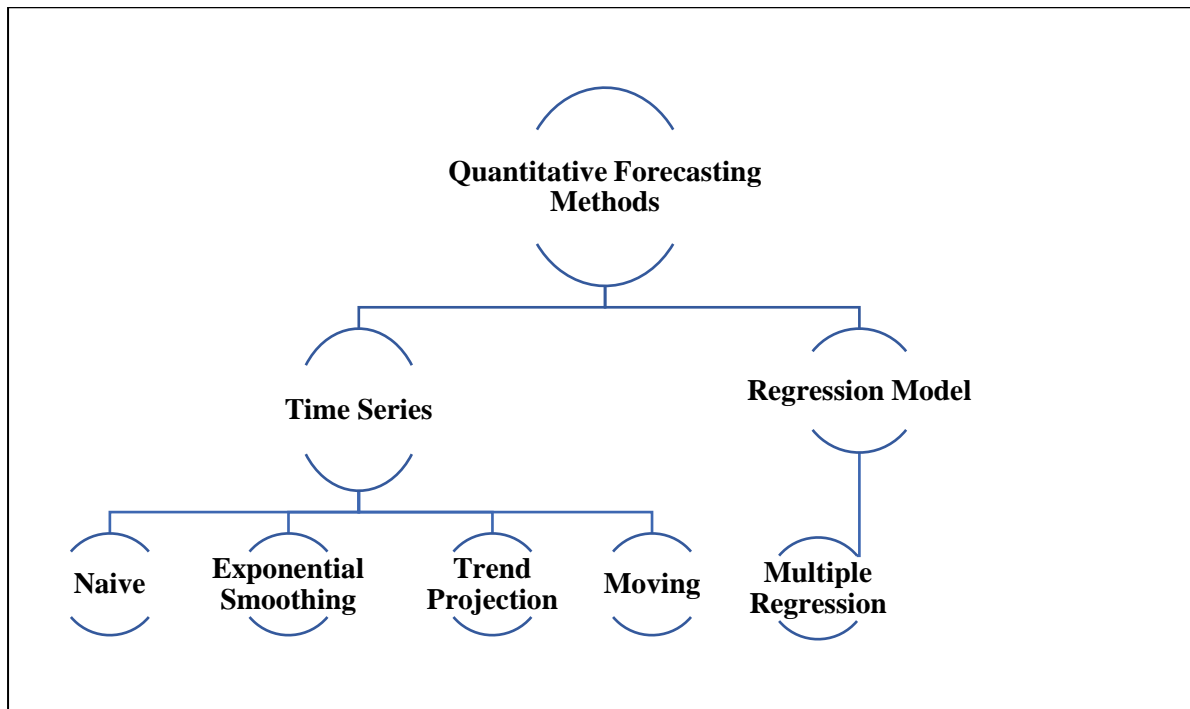


Figure 13-Quantitative forecasting classification

4.1 Time Series Forecasting Model.

A time series [28, 29]forecasting is defined as the time period understood between the previous data is given as input to the system and predicted values refer to the future (a day ahead) or a time series is a series of data point indexed in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time thus it is a sequence of discrete-time data modelled through a stochastic process $y(t)$, i.e. a sequence of random variables is taken from the known data correlated later to obtain forecasting future variable at a time (t) estimating $y(t+h)$ using only information available at the time (t) .

4.1.1 Renewable Production Forecasting Model.

Looking in the electricity production forecast there is currently consent in the literature view a significant to the time series (TS). In terms of TS based in the duration of the time-period it future categorised as Short, medium, and long-term forecast. their respective brief explanation is shortlisted as

Short-term-predicting the future value for the period that can range from few minutes to few days and is of prime importance in day to day market operations. Nowadays short-term technique is fast-moving with benefit to the producer to maximizing the profits. This in retain increases buyers' agents motivated by the benefit from this type of forecast. In short term forecasting with lead time measured is often considered as an isolated class of forecasts

Medium-term-predicting the future values for the period that can range a few days to a few months. Ordinarily, this type of prediction forecast is favoured for balance sheet calculation, risk management and derivatives pricing. As this type of modelling has lengthy-standing tradition inflow finance solution is observed.

Long-term-predicting the future value for the period that can range from months to years This type of prediction forecast is intensified on investment profitability analysis and planning like as determining the future sites or fuel sources[9, 30, 31]

4.1.2 Persistence Forecasting Model:

Model widely used in the short-term forecast of a wind speed or production power. This method assumes that the wind or energy, at a given time future, it will be the same as the value measured at the moment the forecast is made. The forecast corresponds to the average of the most recent N observations in a series X as if you can see in the equation below

$$X_t = \frac{1}{N} \sum_{i=1}^N X_{t-i}$$

This method is simple since the forecast you make is only based on the latest series of values. So it can be used in case the available data, the forecast, insufficient and should only be used for short-term forecasts (1-6 hours).[9]

4.2 Naive Forecasting Methods

This Naive forecast [32] technique basically looking at what happened in the previous period actual are used as future period forecast without modifying them or pursue to establish causal factors. since it is a basic level of forecasting it cannot be that precise efficiency in predicting energy production. This method has some advantage like cost-effective and provides a benchmark against most sophisticated models. This method is quite good while it comes to the economic and financial sectors

4.3 Moving Forecasting Methods

Moving Forecast [32] method is defined as taking the average for historical data of last three periods the next period will probably be the same as those of move that average long whether it can vary by three periods, six periods, or twelve periods

4.4 Exponential smoothing Forecasting Methods

Exponential smoothing [32] practice implement forecasts using weighted averages of the historical past data values and respective forecast error. They are frequently used in inventory regulation systems where the number of things is to be forecast. This technique is low cost which may be a fundamental concern. The simplest exponential smoothing method (SES) reasonable for data with no trend or seasonal patterns.

4.5 Trend Projection Forecasting Methods

Trend projection[33, 34] or least square method is the classical method of business forecasting which is anxious with the flow of fluctuating values along the time. In this method lack, a huge amount of reliable data is required for forecasting. In inclusion, this method is based on assumption is dependent on past trends historical data to predict future trend data in the same manner and to the same extent as it was repeated in the past while determining the variable magnitude and direction. This trend projection method is applied to the long time-series data. This approach mostly used in business and big project prediction. Choosing the right techniques for forecasting along with the quality of information which is been used as an input

4.6 Regression Model

Regression technique [33] is a supervised learning model and analysis is a form of predictive modelling technique which investigates the relationship between a dependent and independent variable. In regression, we can find proper dependencies between variables. It is used to find out how best the variables are related. Regression helps in estimating the relationship between two or more variables. Mainly regression analysis is used to understand the relationship between the dependent and independent variable, that is, what will be the effect on the dependent variable if we change the independent variable. Regression analysis provides a variety of models to perform analysis, for example, linear regression, ordinary least squares, polynomial regression, generalized linear model etc. Mainly, a regression can be linear or multivariate.

4.6.1 Linear Regression

Linear regression[33, 35, 36] is primitive and frequently used for predictive analysis. The overall intention is to establish the relationship between two variables by a linear equation to observed data one variable is considered to be an explanatory variable, and the other is considered to be dependent variable The case of one explanatory variable is known as simple linear regression if the number of explanatory is more than one than that type of process is known as multiple linear regression. Best line to fit two variables is searched so that using one variable we can predict the other. We can visualize it as functional dependencies between two variables, that is, how much the value of the independent variable influences the value of the dependent one.

In this thesis Linear regression is the method used for modelling the relationship between a scalar dependent variable labelled by taking the historical data and predicting the

future here is p_h and one or more interpretive (predictor) variables denoted as \hat{p}_h . The case in which one interpretive or predictor variable is presently called simple linear regression. In linear regression, data is assumed to follow the straight-line relationship and unknown model parameters, that is, the response variable is estimated from the data using the predictor variables. Usually, linear regression is termed as a model in which the conditional mean or value of p_h (called a response variable) for the known value of \hat{p}_h (called a predictor variable), is an affine function (a function having linear function along with constant variable and which contains a graph which has straight line) of \hat{p}_h . i.e., for example, for modelling n data points it requires one independent or predictor variable: \hat{p}_h one dependent or response variable p_h and two parameters, α and β :

$$p_h = \alpha + \beta \hat{p}_h + \varepsilon$$

For $h=1,2,3 \dots n$

where, the variance of p_h is assumed to be constant, and α and β are regression coefficients specifying p_h the intercept and slope of the line, respectively, ε is an error term indicating that for most of the real-world situations, the (p_h, \hat{p}_h) points are not organized exactly in a straight line. More precisely, ε indicates the difference between the actual value and the predicted value.

advantage of linear regression models is easy to understand interpretation on a modular level This is one of the main reasons why the linear model and all similar models are so widespread in academic fields such as medicine, sociology, psychology, and many other quantitative research fields. For example, in the medical field, it is not only important to predict the clinical outcome of a patient, but also to quantify the influence of the drug and at the same time take sex, age, and other features into account in an interpretable way. The linear regression model forces the prediction to be a linear combination of features, which is both its greatest strength and its greatest limitation. Linearity leads to interpretable models. Linear effects are easy to quantify and describe. They are additive, so it is easy to separate the effects. If you suspect feature interactions or a nonlinear association of a feature with the target value, you can add interaction terms or use regression splines.

4.6.2 Multivariate/Multiple Linear Regression

Multiple variables[33] and their correlations and dependencies are studied i.e. data is fit into multiple dimension surfaces to find out the interesting patterns using which we can predict for the future. In general, the model for multivariate linear regression can be given as:

$$p_h = \beta_0 + \beta_1 \hat{p}_1 + \beta_2 \hat{p}_2 + \dots + \beta_n \hat{p}_n + \varepsilon$$

Where,

β_0 is the intercept

β_1 is the parameter associated with \hat{p}_1 (slope)

h represents the number of independent variables $\hat{p}_1 - \hat{p}_h$ independent variables

p_h dependent variable

ε error term

For a given data sample where $p_h = \{ p_1 p_2 p_3 p_4 \dots p_h \}$ and $\hat{p}_h = \{ \hat{p}_1 \hat{p}_2 \hat{p}_3 \hat{p}_4 \dots \hat{p}_h \}$, equation above can be modified as:

$$p_h - \beta_0 + \beta_1 \hat{p}_{n1} + \beta_2 \hat{p}_{n2} + \dots + \beta_h \hat{p}_{nh} + \varepsilon_n$$

Equation above can be simplified and written as,

$$p_h = \beta \hat{p}_h + \varepsilon_n$$

Where

$$\beta = [\beta_0 \beta_1 \beta_2 \dots \beta_h]$$

$$\varepsilon = [\varepsilon_1 \varepsilon_2 \varepsilon_3 \dots \varepsilon_n]$$

The parameter represented by β in the above equation is calculated by using least square estimates so that sum of squared of error is minimized. To compute the coefficient estimates Quadratic Decomposition (QR) method is used. While computing least squares, we calculate the coefficients $\beta_0 \beta_1 \beta_2 \dots \beta_h$ so that the sum of squared errors i.e. SSE is minimum

$$SSE = \sum_{h=1}^n \left(p_h - \beta \hat{p}_h \right)^2$$

Changing feature \hat{p}_h from the reference category to the other category increases the prediction for p_h by β_h when all other features remain fixed. Another important measurement for interpreting linear models is the R-squared measurement. R-squared tells you how much of the total variance of your target outcome is explained by the model. The higher R-squared, the better your model explains the data. The formula for calculating R-squared is $R^2 = 1 - SSE / SST$

$$R^2 = 1 - \frac{\sum_{h=1}^N \left(p_h - \hat{p}_h \right)^2}{\sum_{h=1}^N \left(\hat{p}_h \right)^2}$$

4.7 Artificial neural networks (ANN)

Neural Networks (NN) [9, 37] are data-driven machine learning algorithms which are similar to the smoothing algorithm they learn patterns from the data sets but it also a little bit like regression. NN was designed for capturing a relationship between inputs and output variable using cross-sectional data which have an interesting hybrid. neural networks are also used for forecasting in time series even though they were designed originally just for cross-sectional data and but now a day NN most frequently used method for forecasting

The machine learning algorithms Neural Networks (NN) first become popular in the year 1990s recent period of time it again gained momentum by energy companies and also biggest social network like, Facebook that uses a version of networks called deep learning among all the machine learning algorithms neural networks are the most prominent ones that are used in forecasting time series there actually a book called Neural Networks (NN) in business forecasting by Peter Zend author declare something interesting towards the end that "The development of artificial neural network is still an art rather than a science" there are three types of application in which neural networks forecasting has been used are tourism like forecasting the number of tourists or hotel stays in a country, financial trading and energy sectors including power usage especially renewable energy

Artificial neural networks [9] were developed by the psychiatrist and neuroanatomist Warren McCulloch and mathematician Walter Pitts in 1943. They described a calculation logic of neural networks, which unifies the studies of neurophysiology and logic mathematics. neural networks are highly interconnected simple processing units designed in a way to model how the human brain performs a particular task. Each of those units, also called neurons, forms a weighted sum of its inputs neurons are fundamental nerve cells in the central nervous system, represented by in Figure 14 below. Each person has more than one hundred billion neurons in their system. The construction of a neuron comprises four main parts:

- a) Cellular Body: Pyramidal or spherical in shape, it contains the nucleus inside. This has the task of carrying out the biochemical transformations necessary for the synthesis of molecules, enzymes.
- b) Dendrites: Arborescence around the cell body, formed by thin branches. Are the main signal detectors that reach the neuron, from the axons of the neighbouring nerve cells.
- c) Axon: Nerve fiber branched at the end that communicates with another neuron. This long filament transmits signals emitted by the neuron.
- d) Synapses: It is the connection of the axon of one neuron to the dendrite of another neuron. It is synaptic interface transfers the signals emitted by the neuron to another.

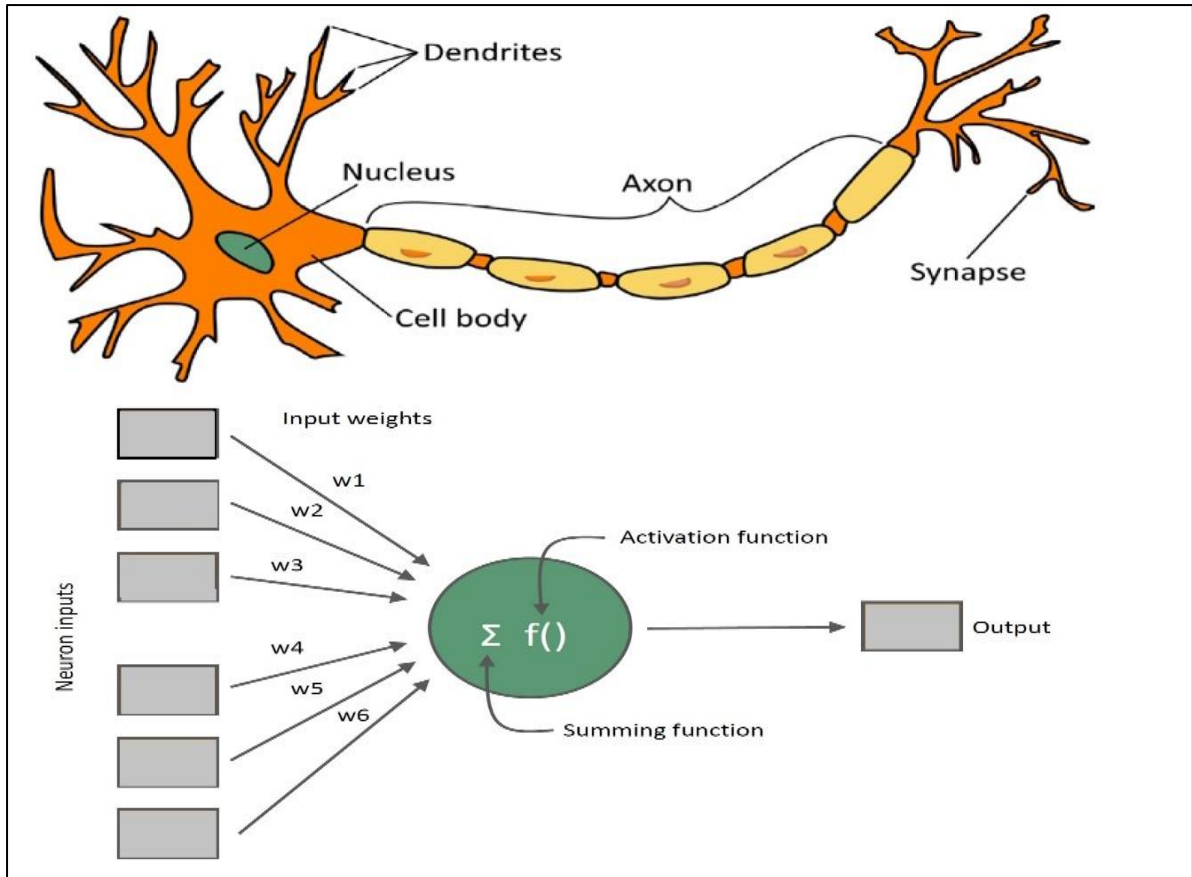


Figure 14-Representation of the four main parts of two biological neurons

Feedforward Neural Networks

Feedforward neural [38] networks allow only unidirectional signal flow. Furthermore, most feedforward neural networks are organized in layers and this architecture is often known as MLP (multilayer perceptron). Multilayer perceptron is the best known and most widely used kind of neural network. networks with interconnections that do not form any loops are called feedforward. recurrent or non-feedforward networks in which there are one or more loops of interconnections are used for some kinds of applications. The units are organized in a way that defines the network architecture. In feedforward networks, units are often arranged in layers: an input layer, one or more hidden layers and an output layer[39]. The units in each layer may share the same inputs but are not connected to each other. Typically, the units in the input layer serve only for transferring the input pattern to the rest of the network, without any processing. The information is processed by the units in the hidden and output layers.

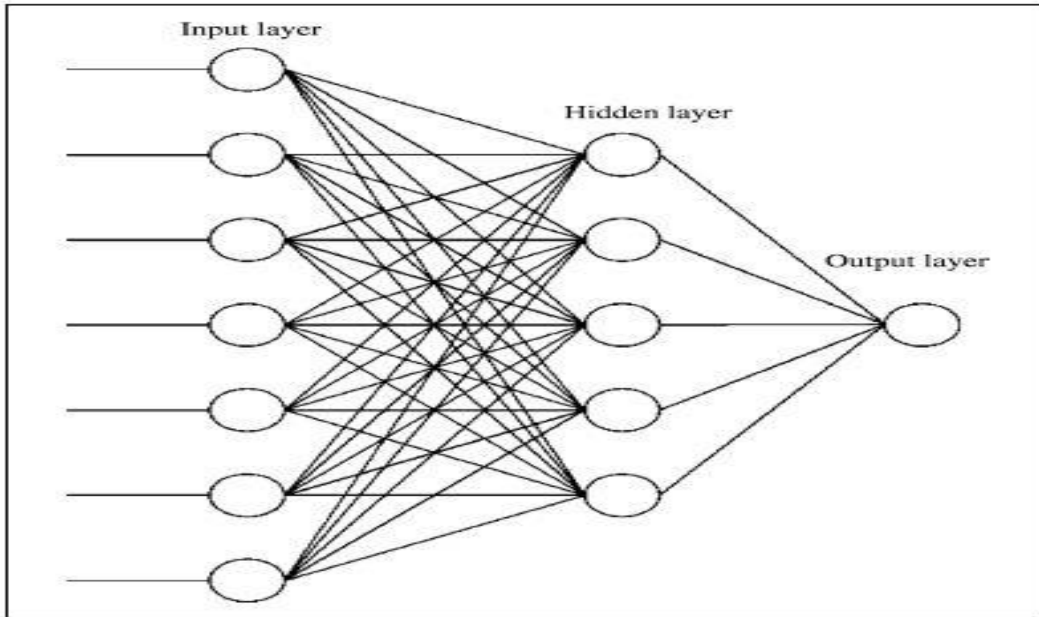


Figure 15-three-layered feedforward neural network model[39]

The imitation learning algorithms, most prevailing use in feed-forward ANN are based on performing analytical gradient plunge on the error. Backpropagation is frequently applied to multiple layers of neurons and works by calculating the overall error rate of an artificial neural network. The output layer is then analysed to see the contribution of each of the neurons to that error. The neurons weights and threshold values are then adjusted, according to how much each neuron subsidize to the error, to minimize the error in the next iteration. To customize its output, several backpropagation learning algorithms use the momentum parameter. Momentum can be useful to prevent a training algorithm from getting trapped in a local minimum and determines how much influence the previous learning iteration will have on the current iteration.[40]

In this study, it was used the learning Levenberg-Marquardt algorithm (Levenberg, 1944 and Marquardt, 1963) that works as a training algorithm with the capabilities of the pruning methodologies. Pruning is a process of examining a solution network, determining which units are not necessary to the solution and removing those units. By doing the artificial neural network prune the model achieved has reduced complexity and the computational effort to run it, especially in real-time, is reduced too. While backpropagation is the steepest descent algorithm, the Levenberg-Marquardt algorithm is a variation of newton's method. The advantage of gauss-newton over the standard newton's method is that it does not require the calculation of second-order derivatives. The Levenberg-Marquardt algorithm trains an ANN faster (10–100 times) than the conventional backpropagation algorithms. The Levenberg-Marquardt algorithm, which consists of finding the update given by:[40]

$$\Delta x = -[J^T(x)J(x) + \mu I]^{-1} J^T(x)e(x)$$

The artificial neuron concept[9], Figure 16, was proposed to simulate a neuron real biological, in which a logical-mathematical structure seeks to simulate its behaviour and functions. This consists of inputs (X_m), such as the synapses of a biological neuron, which are multiplied by the respective weights (W_{km}), representing the signal strength, and finally are

computed by mathematical functions that determine the neuron activation and its polarization, sum function (Σ) and transfer function (T). In addition to these, there is also the computer function of the output (j). Thus, based on the principle established by the authors, the artificial neuron model has the following key elements[9].

- I. Inputs: they constitute the vector $x = [x_1, x_2, x_3, \dots, x_m]^T$, whether of real nature or binary.
- II. Connection weights: positive weights $w = [w_1, w_2, w_3, \dots, w_m]$ in the connections excitations and negatives in inhibitor connections, the weight being null when there is no connection with the respective processing element.
- III. Sum / polarity function: activation threshold of the processing element, equivalent to an entry $x_0 = 1$ with the weight $w_0 = -b_k$
- IV. Transfer function: condition $v = (x, w)$ of excitation of the processing according to the function between the inputs and the weights, $j(v)$ the relation that defines the behaviour of the processing element according to its total internal state
- V. Output: response $y = j(x, w)$ of the neuron resulting from internal processing under excitation of weighted inputs.

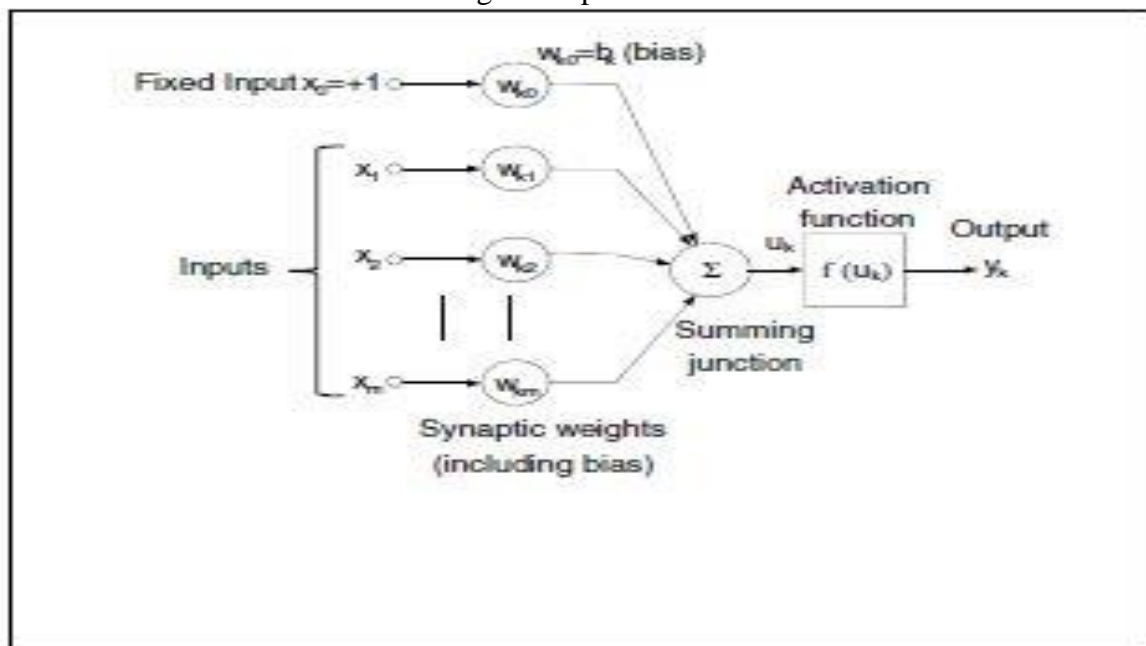


Figure 16-Representation of an artificial neural network.

Based on the thesis of Jessica Lisandra Chaves [40] in a neural network, the inputs excite several neurons that make up a layer: with m elements, there will be m exits, for $k = 1, 2, 3, \dots, m$ forming the output vector $y = [y_1, y_2, y_3, \dots, y_m]$ Each neuron k receives the input signals x_i for $i = 1, 2, 3, \dots, m$ as the weights w_{ki}

Thus, each neuronal network processing element performs the operation imposed by the transfer function, depending on its internal state. In the simplest case, the state total internal corresponds to the linear combination of the inputs with the weights.

$$v_k = w_{k0}x_0 + w_{k1}x_1 + w_{k2}x_2 + \dots + w_{km}x_m$$

when $x_0 = 1$ is assumed then $v_k = \sum_{i=1}^m w_{ki}x_i + b_k$ where the polarity weight b_k translates the original displacement of the activation function and hence the element polarity designation processing.

So, therefore output will be:

$$y_k = \sum_{i=1}^m w_{ki}x_i$$

Or simplifies as

$$y_k = j(w_{ki}x)$$

Where the row vector is w_k and the column vector x [9]

Artificial neural networks are composed of one or more inputs, an output and a or more hidden layers, as seen in the following Figure 17. Each neuron is connected from the input layer to a neuron from the hidden layer, and the hidden layer of each node is connected to a neurone in the output layer[9].

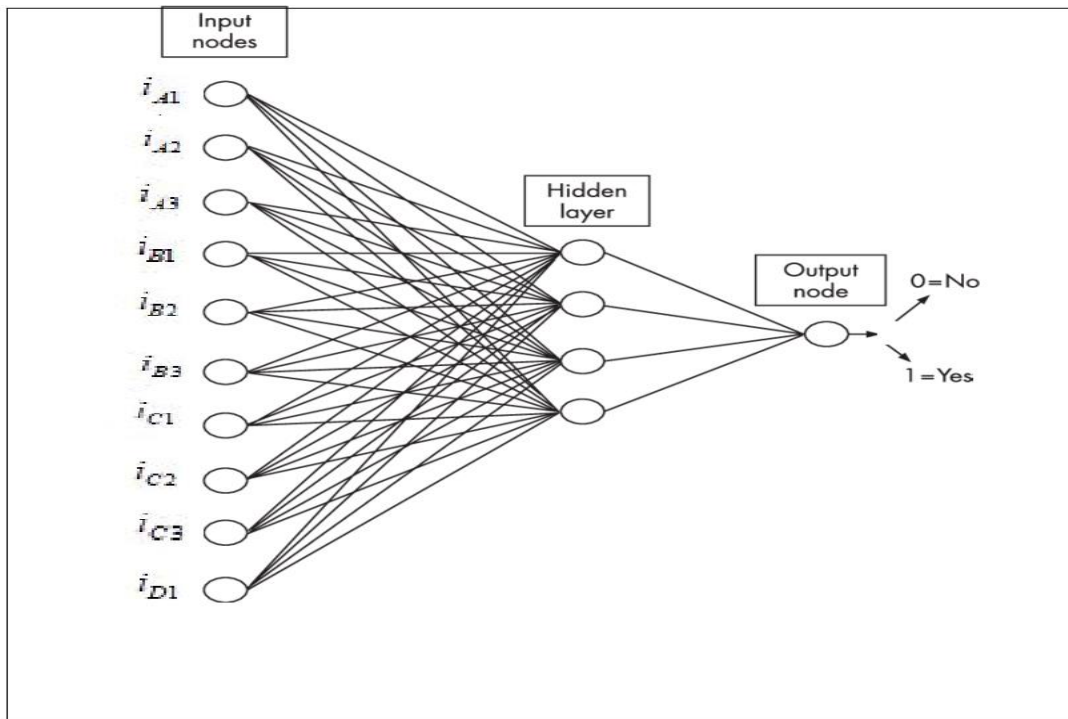


Figure 17-Schematic representation of an artificial neuronal network

5. Forecast Error Measures

Accuracy

In an appraisal of forecasting methods, it is imperative to measure the differences between the possible opportunities. However, it is not obvious which method is the most suitable method to measure forecast errors. Different measures have their advocates. Gardner (2006) compared with five studies approaches and some of its variants. He entrenched that the performance depended on the error measures and the type of data. Teunter and Duncan (2009)[41] discussed that conflicting results may depend on accuracy measures that are not relevant. To reduce the forecast errors to only one measure does not seem possible or even desirable. The different types of forecast errors have different dimensions that make it impossible to reduce them to only one dimension without sacrificing information or data. Behind each measurement lie different methods to calculate the errors and a possibility to obtain complementary information.

To find which model is good and also to check data is helpful in predictions, we will find the prediction error using various available error measurements methods, such as, MSE, RMSE, MAE and MAPE.[33] The forecast error is mainly defined as the difference between the actual value and predicted value. If p_h is the actual value and \hat{p}_h is predicted value then the residual or error term is given as:

$$RE = p_h - \hat{p}_h$$

where, h indicates the number of data points. While forecasting, error measures play an important role, it helps to attune the forecasting models. Refining the forecasting model as per error measures helps in increasing the accuracy of predictions, which is the main goal of predictive analytics.

Definitions for different error measures are given below:

5.1 Mean Squared Error (MSE):

In this case, prediction errors or residuals are squared and then their average is calculated this is known as a mean squared error[33]. A prevailing measure for forecasting errors and its variance is Mean Squared Error (MSE). MSE is related to the standard deviation of forecast errors and is therefore, a relevant error for mathematical operations. In the equation, the assumption is that the mean forecast error deviates from zero and therefore is the mean error present which can be compared to the standard deviation. If instead the mean error is assumed to be equal to zero, the term for mean forecast error is deleted. However, due to the squared function, MSE is more delicate to outliers and errors smaller than one. (Montgomery et al, 1990)[41]

$$MSE = \frac{1}{N} \sum_{h=1}^N \left| p_h - \hat{p}_h \right|^2$$

where n is the total number of data points.

5.2 Root Mean Squared Error (RMSE):

Since the errors were squared to calculate MSE we will take root to get a clear understanding of error. So, the root of the average of the sum of squared error is known as root mean squared error[33].

$$RMSE = \sqrt{\frac{1}{N} \sum_{h=1}^N \left| p_h - \hat{p}_h \right|^2}$$

5.3 Mean Absolute Error (MAE):

The average of the absolute values of the prediction errors or residuals is known as the mean absolute error.

$$MAE = \frac{1}{N} \sum_{h=1}^N \left| p_h - \hat{p}_h \right|$$

5.4 Mean Absolute Percent Error (MAPE):

The average of percent change in the absolute values of the forecasting errors (where percent change means the division of prediction error by actual data) is known as the mean absolute percent error.

$$MAPE = \frac{100}{N} \sum_{h=1}^N \left| \frac{p_h - \hat{p}_h}{p_h} \right|$$

Model having minimum prediction errors are good for prediction cause less the error more will be the accuracy. In the further sections results of experiments will be analysed and discussed using RMSE and MAPE as error measure.

6. Study the Attributes of Energy Production

6.1 Introduction

In arrange to accomplish good production forecasting it is exceptionally critical to choose exceedingly impact variables of the factors related to each technology. It is fundamental to choose the good input variable for each model. Since there are numerous numbers of historical and meteorology variables from which good input variables are chosen. However, few of the choice has been made based on presumptions and a few of the choices were made by a human, so there is conceivable to induce the least blunder. To overcome this sort of blunders pre-processing is escalation imperative when it comes to the choice and organizing the input variables. By this way, the variables were collected from the diverse platform and organized based on technologies and the components impact the factors for that specific technology. This whole segment bargains with study of attributes and analysis of the information variables.[10] .

6.2 Data Collocation and Processing

Based on the detailed study appears that distinctive data source can be utilized to predict production forecast of these information sources have appeared to be valuable to predict production forecast and so it is fundamental to portray for each information source it is used

Abundant of past research has been done on meteorological data variable. MERRA-2 meteorological variables are noteworthy for day-ahead forecast since they covering a skyline of 24 to 144 hours ahead back. The data variables of meteorological were collected from MERRA-2 webpage platform. All the data variables measured wind speed at 10-meter statue from the ground level, wind direction at 10-meter tallness from the ground level, Temperature at 2-meter over ground level and relative humidity at 2-meter over the ground level which portrays the condition of the climate counting considerable factors such as wind speed, wind direction, temperature, irradiations, pluviosity and relative-humidity are indispensable which inevitably move forward to the accuracy of prediction beside the quality of meteorological information factors were taken from MERRA-2

Within the to begin with stage, the information accessible on the energy production and utilization for Portugal, the source gives the data of 12 months and as it was permitted to access by Excel files secured the time interval based on the technologies which are shown below

Table 4-Technology and Time Interval

Technology	Time Interval
Wind Farm	01/01/2018
Solar Power Plant	01/01/2018
Hydro Power Plant	01/01/2019

This information data and utilization were accessible on the service platform Kisense energy Optimizer enlisted in Portugal regarding the approach of energy production. Utilization, import/export data sum of energy produced by each sort of technology within the periods of one hour for the whole year (Table 4). This information data stage altogether moves forward the adequacy of the forecasting models to built, with regard to short-term forecast (at slightest one day-one week)[10]. The same sort of information data related to each technology.it was attainable to secure these values on the stage Kisense [42]

Nearly comparable to the production information data stage Kisense there is doable Meteorology platform to access the weather data. Since there was a numerous of website related to these divisions it was quite difficult to require in thought. Based on the need of the proposal objective the finest alternative was MERRA-2[43]. This web platform holds the benefit NASA(National Aeronautics and Space Administration) which leads to more exactness towards the information data. In this manner, this accuracy of information would be supportive to accomplish productivity forecasting results.

This web platform information is presented under certain priority it is shortlisted below

Table 5 Weather Data specification

Modern-Era Retrospective Analysis for Research and Applications (MERRA)-version 2 Provider National Aeronautics and Space Administration (NASA) / Goddard Space Flight Center More information at- http://gmao.gsfc.nasa.gov/reanalysis/MERRA-2	
Site latitude (positive means North)	39
Site longitude (positive means East)	-9
Site altitude (m)	71
Date beginning	01-01-2018
Date end	31-12-2018
Time reference (hour)	UT
Summarization (period of integration)	Hour (h)
Date - Instant The date or instant given	for each value corresponds to the end of the summarization
Columns:	
Temperature (K)	Temperature at 2 m above ground
Relative humidity (%)	Relative humidity at 2 m above ground
Pressure (hPa)	Pressure at ground level
Wind speed (m/s)	Wind speed at 10 m above ground
Wind direction (deg)	Wind direction at 10 m above ground (0 means from North 90 from East...)
Rainfall (kg/m2)	Rainfall (= rain depth in mm)
Snowfall (kg/m2)	Snowfall
Snow depth (m)	Snow depth
Short-wave irradiation (Wh/m2)	Surface incoming shortwave irradiation (broadband)

Utilization, import/export of Meteorology data variable for each sort of technology within the particular time periods. Each technology imported data period of each hour for the whole year This driven to a diverse approach and from there to an adjustment of the data that was collected from both data sources so that the data may be treated together. This adaptation included data gathering based on the technology which arranges to create a single model.

There are few approaches in analyzing of data variables some time recently applying to chosen technique and in arrange to summarize their primary characteristics based on the regulating qualitative measures such as cruel (blunder/error risk management) and relationship. The labels of a few attributes are shortlisted underneath.[43].

Table 6 Precipitation data lables

Temperature (T)	Avg Temperature (AvgT)
	Min Temperature (MinT)
	Max Temperature (MaxT)
Wind speed (WS)	Avg Wind speed (AvgWS)
	Min Wind speed (MinWS)
	Max Wind speed (MaxWS)
Wind direction (WD)	Avg Wind direction (AvgWD)
	Min Wind direction (MinWD)
	Max Wind direction (MaxWD)
Rainfall Snowfall/Pluviosity (P)	Avg Pluviosity (AvgP)
	Min Pluviosity (MinP)
	Max Pluviosity (MaxP)
Short-wave irradiation/ irradiation(I)	Avg irradiation(Avg I)
	Min irradiation(Min I)
	Max irradiation(Max I)
Relative Humidity(H)	Avg Humidity(AvgH)
	Min Humidity(MinH)
	Max Humidity(MaxH)

Afterwards, arrange the average of all these variables for each day in arrange to get average esteem agent of the Iberian Peninsula. A test of the data treated with regard to meteorological data can be counselled within the annex. Other meteorological data were taken into thought was solar radiation. This data was gotten from site Photovoltaic Geographical Information System (PGIS) [44]. the average information for each month, speaking to verifiable values from 2005 to 2016 a long time, be that as it may, they down to business data and agent of reality, this website being utilized to form calculations for the establishment of photovoltaic panels Annex presents the data considered this information.

This practical information data collected for each and each point went through precipitation information, such as.

Table 7 Precipitation photovoltaic data lables

Avg-H _h	Irradiation on horizontal plane (kWh/m2/mo)
Avg-H _{i_opt}	Irradiation on optimally inclined plane (kWh/m2/mo)
Avg-H _{i_m}	Irradiation on plane at angle (kWh/m2/mo)
Avg-H _{b(n)_m}	Monthly beam (direct) irradiation on a plane always normal to sun rays (kWh/m2/mo)

6.3. Correlation Analysis

In arranging upsurge the accuracy of the forecast. A nitty gritty examination of variable relationship is carried out from the collected data. A correlation between factors shows that as one variable changes in esteem, the other variable tends to alter in a particular direction. Understanding that relationship is valuable since we will utilize the esteem of one variable to predict the esteem of the other variable [10, 45].

6.3.1 Pearson Correlation Coefficient

The Pearson correlation coefficient (PCC) too alluded as Pearson's or, the Pearson product-moment correlation coefficient (PPMCC), or the bivariate correlation. This strategy measurement that measures the linear correlation between two variables X and Y. It includes esteem of run between +1 and -1. The esteem of run falls on +1 at that point it is spoken to as Positive linear correlation, esteem falls on the extent it is represented as no linear correlation and esteem falls on the run -1 it is represented as Negative linear correlation[45-47].

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

The best degree of correlation which highly depends on the values falling range of correlation +1 to -1

Perfect Correlation:

The correlation values are near to $\frac{+}{-}1$ or 100%, then it represents to be a perfect correlation as one variable upsurge, the other variable tends to also upsurge (if positive) or decrease (if negative) [45].

High Correlation:

In this dissertation in some case the correlation values lie between $\frac{+}{-}0.50$ and $\frac{+}{-}1$, or 50% and 100% then it is represented or said to be a strong correlation

Moderate Correlation:

In this dissertation in some case the correlation values lie between $\frac{+}{-}0.30$ and $\frac{+}{-}0.49$ or 30% and 49%, then it is represented or said to be a Medium correlation[45]

Low Correlation:

In this dissertation in some case the correlation values lie beneath $\frac{+}{-}0.29$ or 29%, then it is said to be a small correlation

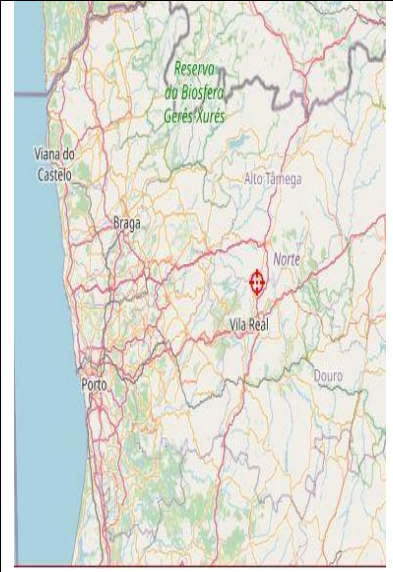

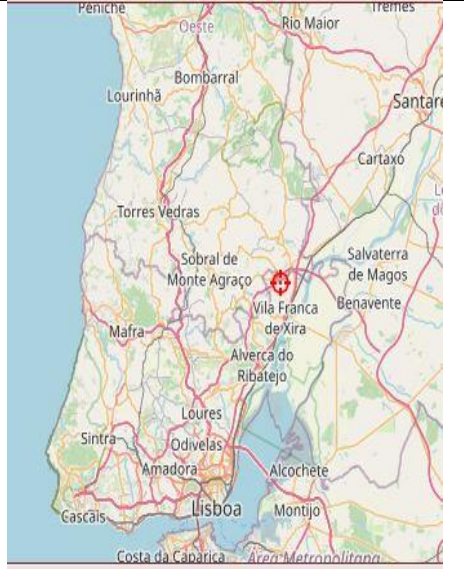
No Correlation:

When the value is zero[45].

6.4 Meteorological Data

The meteorological information variable from MERE-2 has been utilized for the prediction of each technologies production have been given by the National Aeronautics and Space Administration (NASA). NASA has utilized the meteorological model to appraise the weather estimate for diverse geographical latitude and longitude facilitates. The facilitates from a framework covering Portugal can be seen within the figure based on the technologies. The figure gives data of the area of each technologies production forecasting by the development of assorted input variables which is chosen based on the correlation relationship.

Table 8 Location of each technologies

Hydro	Solar	Wind
		

The gotten data that were treated and tested will be displayed in relation to Production and Meteorological data. A clear relationship between the variables is illustrated to the production at slightest on an hourly base and daily which is spoken to in this destination. The gotten data set there is an awesome differing quality of the sort of variables, that were classified concurring to their characteristics.[10].

6.4.1 Historical Variables based on Technology

Ki-sense is an energy production platform for each technology provided by the company Simples Energia permitted to access the data variables which were related to this dissertation. The data variable was collected in the excel spreadsheet. The collected data variable comprises of that specific day data were represented with hourly information. For predicting future production forecasting, some influencing factor should be used that may help to achieve the accuracy of the forecast prediction of production. Concurring to detail consider of meteorological data variable few influence factor such as the day before which is symbolized by D-1, the previous day before which is symbolized by D-2, three days before which is symbolized by D-3 and so on until D-6.

The notation used as D-1, D-2, D-3, D-4, D-5 and D-6 with signify time period of difference is 24 hours backwards (i.e.) basically The documentation utilized as D-1, for the case, demonstrates the esteem of the alluded variable, 24 hours before the time for which need to forecast (day-1). Moreover, D-2 alludes to 48h before similarly D-3 72 hours before, D-4 is 96 hours before, D-5 is 120 hours before and D-6 is 144 hours before. Here D is referred to as historical data variables of the production. This reference is shortlisted below the It is a fundamental accentuation to carried out the correlation of historical variables notation as “D” with the actual production which is notation as “A”.

The most reason of correlation is between variables imply that as the esteem of one variable varies the other variables doable to varies in a particular direction. By compassionate, the relationship is valuable since with the utilize of this approach values of one variable to predict the esteem of the other variables. This handle is been executed for all the technologies in arrange to decide the correlation of the historical variables. After determining the correlation, as it were a better rate of correlation, will be assumed as input for the forecasting model which is able to be seen in afterwards sections. Table 9.

Table 9 Example how the historical production (Daily) data are regulated

Date	Hour	Day of the week	Month	Actual Production (A)	D-1	D-2	D-3	D-4	D-5	D-6
01-01-2019	23:00	1	1	X						
02-01-2019	23:00	1	1	X	X					
03-01-2019	23:00	1	1	X	X	X				
04-01-2019	23:00	1	1	X	X	X	X			
05-01-2019	23:00	1	1	X	X	X	X	X		
06-01-2019	23:00	2	1	X	X	X	X	X	X	
07-01-2019	23:00	2	1	X	X	X	X	X	X	X
08-01-2019	23:00	2	1	X	X	X	X	X	X	X
09-01-2019	23:00	2	1	X	X	X	X	X	X	X
10-01-2019	23:00	2	1	X	X	X	X	X	X	X
11-01-2019	23:00	2	1	X	X	X	X	X	X	X

It basically from where D-1 data variables decide its essential from actual production variable with notation utilized is “A”. which variables are bolstered to the historical variable but it is essential to check the data variables were put on that particular day with regard to time-period. This prepare is taken after for rest all verifiable information factors which are D-2 to D-6. These by this way information variable were collected and organized for hourly models so also information was collected and organized for day by day models were too built.

6.4.3 Correlation of Historical Variables for All Technology

It is a fundamental accentuation to carried out the correlation of historical variables notation as “D” with the actual production which is notation as “A”. the most reason of correlation is between variables imply that as the esteem of one variable varies the other variables doable to varies in a particular direction. By compassionate, the relationship is valuable since with the utilize of this approach values of one variable to predict the esteem of the other variables. This handle is been executed for all the technologies in arrange to decide the correlation of the historical variables. After determining the correlation, as it were a better rate of correlation, will be assumed as input for the forecasting model which is able to be seen in afterwards sections.

Table 10. Production Correlation with Previous day hourly/daily Production by Technology.

Correlation Coefficients relative to Production Using Pearson’s Coefficient		
Tested Variable	Historical Production reference	Correlation Value in Percentage
Hydro(hourly)	D-1	85.49%
	D-2	78.34%
	D-3	69.49%
	D-4	62.70%
	D-5	59.38%
	D-6	56.41%
Hydro(Daily)	D-1	91.95%
	D-2	82.95%
	D-3	73.73%
	D-4	66.22%
	D-5	62.18%
	D-6	58.90%
Solar(Hourly)	D-1	88.83%
	D-2	87.54%
	D-3	86.94%
	D-4	86.82%
	D-5	86.49%
	D-6	86.26%
Solar(Daily)	D-1	70.84%
	D-2	66.39%
	D-3	64.47%
	D-4	63.73%
	D-5	61.70%
	D-6	62.00%
Wind (Hourly)	D-1	35.82%
	D-2	20.49%
	D-3	13.94%
	D-4	12.45%
	D-5	8.64%
	D-6	4.38%
Wind (Daily)	D-1	47.85%
	D-2	24.52%
	D-3	13.44%
	D-4	11.82%
	D-5	6.74%
	D-6	1.13%

Here are presented some of the most prevalent variables that have been tested for their inclusion in the Production forecasting Models. Initially, the correlation with Production load (Actual Production), this data may be Influential with the view of forecast Models. Likewise, the relationship between some variables along with load consumed, essential to build a forecast model for an hourly and daily base to estimated, serving, and hoping it will improve Production forecast.

6.4.4 Meteorological Variables for Hydro Technology

MERRA-2 is meteorological data variable platform for hydro technology which provide few data variables which were influencing for the forecasting for this hydro technology. This data variable was collected in the excel spreadsheet from MERRA-2 website. The collected data variable was of that specific day with variables hourly of all meteorological data. For predicting the future production some influencing factor would help to achieve the accuracy of the forecast production models.

Based on detail study of meteorological data variable few influence factor such as pluviosity and relative-humidity notation as (avg, min, max) which is represented as P, min(P), max(P) avg(P) for pluviosity and avg(H), min(H), max(H) for relative-humidity. The notation used for this meteorological data variable is used to estimate meteorological historical reference such as D-1, D-2, D-3, D-4, D-5 and D-6 with the time difference of 24 hours so it is basically 24h,48h, 72h, 96h, 120h and 144h respectively. All these variables represent meteorological historical variables of weather data.

It is essentially from where D-1 meteorology data variables are determined it is fundamentally from meteorological data such as P, Avg P, Min P, Max P, Avg H, Min H and Max H variables values which feed meteorological data of previous day data is feed to next day with respect to the specific same time period. This process is repeated for rest all meteorological data i.e. from D-2 to D-6 variables which lead to the formation of historical meteorological data. In this way, data was organized and utilized to develop a production forecast for hydro technology. This is achieved by construction numerous numbers of hourly and daily modes. The input selected to each model is based on the correlation relationship of different variable combination.

In this Hydro Technology, there was very little time eliminated with the values due to no production in that period. This no production term can be due to the need for precipitation in that time of period, or overwhelming maintenance within the same period that could moreover attainable. Comprehension to eliminate this variable could lead to significant towards accomplishing the accuracy of the result. In this dissertation, eliminated variables period for 6/9/2019 to 15/10/2019 this is often as it were applicable on for this technology.

Table 11-Correlation Coefficients relative to Production (Hourly) Vs Meteorological Using Pearson’s Coefficient

Correlation Coefficients relative to Production Vs Metrological Using Pearson’s Coefficient			
Tested Variable	Metrological reference	Metrological reference	Correlation Value in Percentge
Hydro(hourly)	D (Correlation Day (Hourly))	(P)	41.57%
		Min (P)	42.11%
		Max (P)	30.12%
		Avg (H)	61.42%
		Min (H)	64.70%
		Max (H)	32.37%
		D-1 (Correlation Day-1 (Hourly))	Avg (P)
	Min (P)		45.37%
	Max (P)		32.86%
	Avg (H)		64.46%
	Min (H)		69.60%
	Max (H)		31.22%
	D-2 (Correlation Day-2 (Hourly))	Sum (P)	51.84%
		Avg (H)	65.91%
	D-3 (Correlation Day-3 (Hourly))	Sum (P)	55.60%
		Avg (H)	67.22%
	D-4 (Correlation Day-4 (Hourly))	Sum (P)	55.89%
		Avg (H)	68.06%
	D-5 (Correlation Day-5 (Hourly))	Sum (P)	56.08%
		Avg (H)	68.39%
D-6 (Correlation Day-6 (Hourly))	Sum (P)	55.97%	
	Avg (H)	68.36%	
Month (Hourly)		23.28%	

6.4.5 Correlation of Meteorological Variables for Hydro Technology

It is essentially emphasis to carried out the correlation of meteorological historical notation as “sum(P), min(P), max (P), Avg (H), min (H) max (H) and month (m)” variables with actual production variables which is notation as “A”. in order to know the main purpose of correlation first we have to understand the relationship between the variables that is one variable is related to other variables by this logic utilize to predict future forecast production value. The correlation of meteorological variable for hydro technology is shown in

Table 12 refer to the total input which can be added to serve as input to the Neural Network (NN) it favorably output can be attained.

To procure best results some variable were chopped off, despite having a good correlation coefficient. By eliminated or chopped off the variable of the same specification for example, consider Avg(P) for D, D-1, D-2, D-3, D-4, D-5, D-6 values can be similar which may complex the Network to attain the accurate result. To overcome this complexity certain variable was eliminated this eliminated variable was done based on the correlation and for every model, input variable was phased out or eliminated. This is followed on both hourly, daily every variable class-based data for each technology.

Table 12 Correlation Coefficients relative to Production (Daily) Vs Meteorological Using Pearson’s Coefficient

Correlation Coefficients relative to Production Vs Metrological Using Pearson’s Coefficient			
Tested Variable	Metrological reference	Metrological reference	Correlation Value in Percentage
Hydro(Daily)	D (Correlation Day (Daily))	Sum (P)	45.00%
		Min (P)	45.58%
		Max (P)	32.61%
		Avg (H)	66.48%
		Min (H)	70.03%
		Max (H)	35.03%
		D-1 (Correlation Day-1 (Daily))	Avg (P)
	Min (P)		49.10%
	Max (P)		35.56%
	Avg (H)		69.77%
	Min (H)		75.32%
	Max (H)		33.80%
	D-2 (Correlation Day-2 (Daily))	Sum (P)	56.08%
		Avg (H)	71.32%
	D-3 (Correlation Day-3 (Daily))	Sum (P)	60.15%
		Avg (H)	72.73%
	D-4 (Correlation Day-4 (Daily))	Sum (P)	60.45%
		Avg (H)	73.62%
	D-5 (Correlation Day-5 (Daily))	Sum (P)	60.65%
		Avg (H)	73.98%
D-6 (Correlation Day-6 (Daily))	Sum (P)	60.53%	
	Avg (H)	73.94%	
Month (Daily)	M	25.26%	

Initially, from the above section will be briefly confirmable, the entries with the reference “D-1, D-2, D-3, D-4, D-5, D-6” would be used as inputs for the model 1. However, after checking that correlation of data only higher correlation were consider for each type of variable P(D-1), Sum P(D-5), Min P(D-1), Avg H(D-5) and Min H(D-1). For every Input Models, there were eliminated or chopped off variables. This eliminated will take place until the least correlation variable of a class type exist.

6.4.6 Meteorological Variables for Solar Technology

MERRA-2 is meteorological data variable platform for solar technology which provide few data variables which were influencing for the forecasting for this solar technology. This data variable was collected in the excel spreadsheet from MERRA-2 website. The collected data variable was of that specific day with variables hourly of all meteorological data. For predicting the future production some influencing factor would help to achieve the accuracy of the forecast production models.

Based on detail study of meteorological data variable few influence factor such as temperature and irradiation notation as (avg, min, max) which is represented as min(T), max(T) avg(T) for temperature and avg(I), min(I), max(I) for irradiation. The notation used for this meteorological data variable is used to estimate meteorological historical reference such as D-1, D-2, D-3, D-4, D-5 and D-6 with the time difference of 24 hours so it is basically 24h,48h, 72h, 96h, 120h and 144h respectively. All these variables represent meteorological historical variables of weather data

It is essentially from where D-1 meteorology data variables are determined it is fundamentally from meteorological data such as Avg T, Min T, Max T, Avg I, Min I and Max I variables values which feed meteorological data of previous day data is feed to next day with respect to the specific same time period. This process is repeated for rest all meteorological data i.e. from D-2 to D-6 variables which lead to the formation of historical meteorological data. In this way, data was organized and utilized to develop a production forecast for solar technology. This is achieved by construction numerous numbers of hourly and daily modes. The input selected to each model is based on the correlation relationship of different variable combination.

Meteorological data collected for previously spoken and these revealed to have, in general, more influential values in the development the production. Together with these data it is a fundamental accentuation to carried out the correlation of historical variable notation as “D” with the actual production which is notation as “A”. The most reason of correlation is between variables imply that as the value of one variable varies the other variables doable to varies in a particular direction. By compassionate, the relationship is valuable since with the utilize of this approach values of one variable to predict the value of the other variables. This handle is been executed for all the technologies in arrange to decide the correlation of the historical variables. After determining the correlation, as it were a good rate of correlation will be assumed as input for the forecasting model which is able to be seen afterwards section and meteorological data, forecasting models were been created, although another source and hourly values and of daily as in the other cases, will be included the solar radiation values. The correlations of the spoken data can be verified in Table 15

Table 13 Correlation Coefficients relative to Solar Production (Hourly) Vs Meteorological Using Pearson’s Coefficient

Correlation Coefficients relative to Production Vs Metrological Using Pearson’s Coefficient			
Tested Variable	Metrological reference	Metrological reference	Correlation Value in Percentage
Solar (hourly)	D (Correlation Day (hourly))	Avg(T)	20.77%
		Min (T)	17.86%
		Max (T)	22.45%
		Avg (I)	21.52%
		Max (I)	21.03%
	D-1 (Correlation Day-1 (hourly))	Avg(T)	21.10%
		Min (T)	18.69%
		Max (T)	22.06%
		Avg (I)	19.46%
		Max (I)	18.40%
	D-2 (Correlation Day-2 (hourly))	Avg(T)	21.17%
		Max (T)	22.02%
		Avg (I)	19.79%
	D-3 (Correlation Day-3 (hourly))	Avg(T)	21.11%
		Max (T)	21.92%
		Avg (I)	19.83%
	D-4 (Correlation Day-4 (hourly))	Avg(T)	20.94%
		Max (T)	21.67%
		Avg (I)	19.74%
	D-5 (Correlation Day-5 (hourly))	Avg(T)	20.77%
		Max (T)	21.44%
		Avg (I)	19.61%
	D-6 (Correlation Day-6 (hourly))	Avg(T)	20.76%
		Max (T)	21.35%
Avg (I)		19.63%	

6.4.7 Correlation of Meteorological Variables for Solar Technology

It is essentially emphasis to carried out the correlation of meteorological historical notation as “Avg(T), min(T), max (T), Avg (I), min (I) max (I) and month (m)” variables with actual production variables which is notation as “A”. To know the main purpose of correlation first we have to understand the relationship between the variables that is one variable is related to other variables by this logic utilize to predict future forecast production value. The correlation of meteorological variable for hydro technology is shown inTable 14Table 15. Data variable with higher correlation variable is selected as the input variable which will be seen in the upcoming section

Analyzing the values presented in this solar technology, we can see the importance of both solar radiation, temperature, in the evolution of the production. Even the irradiation, shows to have more effect on. These third values will then be considered for the production forecasting model. The values of these Pearson coefficients are shown in Table 14

Table 14 Correlation Coefficients relative to Solar Production Vs Meteorological Irradiation Using Pearson’s Coefficient

Correlation Coefficients relative to Production Vs Metrological Using Pearson’s Coefficient			
Technology	Metrological reference	Metrological reference	Correlation Value in Percetage
Solar	Hourly	Avg-Hh	18.49%
		Avg-Hi_opt	19.23%
		Avg-Hi_m	19.06%
		Avg-Hb(n)_m	19.36%
	Daily	Avg-Hh	62.94%
		Avg-Hi_opt	65.43%
		Avg-Hi_m	64.86%
		Avg-Hb(n)_m	65.90%

Very similar to hourly data variables same steps were been followed daily data variables It is a fundamental accentuation to carried out the correlation of historical variables notation as “D” with the actual production which is notation as “A”. the most reason of correlation is between variables imply that as the esteem of one variable varies the other variables doable to varies in a particular direction. By compassionate, the relationship is valuable since with the utilize of this approach values of one variable to predict the esteem of the other variables. This handle is been executed for all the technologies in arrange to decide the correlation of the historical variables. After determining the correlation, as it were a better rate of correlation, will be assumed as input for the forecasting model which is able to be seen in afterwards sections.

Table 10 refers to historical data variable correlate with actual production values analogous Table 15 refers to correlation coefficients relative to solar daily production along with Meteorological data variable.

Table 15 Correlation Coefficients relative to Solar Production (Daily) Vs Meteorological Using Pearson’s Coefficient

Correlation Coefficients relative to Production Vs Metrological Using Pearson’s Coefficient			
Tested Variable	Metrological reference	Metrological reference	Correlation Value in Percentage
Solar (Daily)	D (Correlation Day (Daily))	Avg(T)	70.64%
		Min (T)	60.75%
		Max (T)	76.34%
		Avg (I)	73.17%
		Max (I)	71.48%
	D-1 (Correlation Day-1 (Daily))	Avg(T)	71.78%
		Min (T)	63.59%
		Max (T)	75.07%
		Avg (I)	66.17%
		Max (I)	62.58%
	D-2 (Correlation Day-2 (Daily))	Avg(T)	72.14%
		Max (T)	75.05%
		Avg (I)	67.42%
	D-3 (Correlation Day-3 (Daily))	Avg(T)	72.27%
		Max (T)	75.04%
		Avg (I)	67.86%
	D-4 (Correlation Day-4 (Daily))	Avg(T)	72.01%
		Max (T)	74.52%
		Avg (I)	67.84%
	D-5 (Correlation Day-5 (Daily))	Avg(T)	71.87%
		Max (T)	74.19%
Avg (I)		67.79%	
D-6 (Correlation Day-6 (Daily))	Avg(T)	71.78%	
	Max (T)	73.80%	
	Avg (I)	67.80%	
Month (Hourly)	M	20.91%	

Initially, from the above section will be briefly confirmable, the entries with the reference “D-1, D-2, D-3” would be used as inputs for the model 1. However, after checking that correlation of data only higher correlation were consider for each type of variable Max of T (D), Avg I (D), Max I (D), Min T (D-1), Avg T (D-4), Avg-Hh, Avg-Hi_opt, Avg-Hi_m, Avg-Hb(n)_m. For every input model, there were eliminated or chopped off variables. This eliminated will take place until the least correlation variable of a class type exist.

6.4.8 Meteorological Variables for Wind technology

MERRA-2 is meteorological data variable platform for wind technology which provide few data variables which were influencing for the forecasting for this wind technology. This data variable was collected in the excel spreadsheet from MERRA-2 website. The collected data variable was of that specific day with variables hourly of all meteorological data. For predicting the future production some influencing factor would help to achieve the accuracy of the forecast production models.

Based on detail study of meteorological data variable few influence factor such as wind speed and wind direction notation as (avg, min, max) which is represented as min(WS), max(WS) avg(WS) for wind speed and avg(WD), min(WD), max(WD) for wind direction. The notation used for this meteorological data variable is used to estimate meteorological historical reference such as D-1, D-2, D-3, D-4, D-5 and D-6 with the time difference of 24 hours so it is basically 24h,48h, 72h, 96h, 120h and 144h respectively. All these variables represent meteorological historical variables of weather data

It is essentially from where D-1 meteorology data variables are determined it is fundamentally from meteorological data such as Avg WS, Min WS, Max WS, Avg WD, Min WD and Max WD variables values which feed meteorological data of previous day data is feed to next day with respect to a specific same time period. This process is repeated for rest all meteorological data i.e. from D-2 to D-6 [31] variables which lead to the formation of historical meteorological data. In this way, data was organized and utilized to develop a production forecast for wind technology. This is achieved by construction numerous numbers of hourly and daily modes. The input selected to each model is based on the correlation relationship of different variable combination.

Meteorological data collected for previously spoken and these revealed to have, in general, more influential values in the development the production. Together with these data It is a fundamental accentuation to carried out the correlation of historical variables notation as “D” with the actual production which is notation as “A”. the most reason of correlation is between variables imply that as the esteem of one variable varies the other variables doable to varies in a particular direction. By compassionate, the relationship is valuable since with the utilize of this approach values of one variable to predict the esteem of the other variables. This handle is been executed for all the technologies in arrange to decide the correlation of the historical variables. After determining the correlation, as it were a better rate of correlation, will be assumed as input for the forecasting model which is able to be seen in afterwards sections.

Table 10 meteorological data, forecasting models were been created, although another source and hourly values and of daily models. The correlations of the spoken data can be verified in Table 16

Table 16 Correlation Coefficients relative to Wind Production (Hourly) Vs Meteorological Using Pearson’s Coefficient

Correlation Coefficients relative to Production Vs Metrological Using Pearson’s Coefficient			
Tested Variable	Metrological reference	Metrological reference	Correlation Value in Percentage
Wind (hourly)	D (Correlation Day (hourly))	Avg(WS)	60.11%
		Min (WS)	58.42%
		Max (WS)	49.64%
		Avg (WD)	31.40%
		Min (WD)	28.11%
		Max (WD)	12.03%
	D-1 (Correlation Day-1 (hourly))	Avg(WS)	33.50%
		Min (WS)	35.01%
		Max (WS)	30.74%
		Avg (WD)	21.84%
		Min (WD)	22.32%
		Max (WD)	10.10%
	D-2 (Correlation Day-2 (hourly))	Avg(WS)	27.38%
		Max (WS)	22.71%
		Avg (WD)	19.99%
	D-3 (Correlation Day-3 (hourly))	Avg(WS)	22.97%
		Max (WS)	18.29%
		Avg (WD)	18.44%
	D-4 (Correlation Day-4 (hourly))	Avg(WS)	21.13%
		Max (WS)	15.59%
		Avg (WD)	16.25%
	D-5 (Correlation Day-5 (hourly))	Avg(WS)	20.11%
		Max (WS)	13.13%
		Avg (WD)	13.38%
D-6 (Correlation Day-6 (hourly))	Avg(WS)	18.12%	
	Max (WS)	10.28%	
	Avg (WD)	12.68%	

6.4.9 Correlation of Meteorological Variables for Wind Technology

It is essentially emphasis to carried out the correlation of meteorological historical notation as “Avg(WS), min(WS), max (WS), Avg (WD), min (WD) max (WD) and month (m)” variables with actual production variables which is notation as “A”. To know the main purpose of correlation first we must understand the relationship between the variables that is one variable is related to other variable by this logic utilize to predict future forecast production value. The correlation of meteorological variable for wind technology is shown in Table 16 Table 17. Data variable with higher correlation variable is selected as an input variable which will be seen in the upcoming section

Very similar to hourly data variables same steps were been followed daily data variables It is a fundamental accentuation to carried out the correlation of historical variables notation as “D” with the actual production which is notation as “A”. the most reason of correlation is between variables imply that as the esteem of one variable varies the other variables doable to varies in a particular direction. By compassionate, the relationship is valuable since with the utilize of this approach values of one variable to predict the esteem of

the other variables. This handle is been executed for all the technologies in arrange to decide the correlation of the historical variables. After determining the correlation, as it were a better rate of correlation, will be assumed as input for the forecasting model which is able to be seen in afterwards sections.

Table 10 refers to historical data variable correlate with actual production values analogous Table 17 refers to correlation coefficients relative to solar daily production along with Meteorological data variable

Table 17 Correlation Coefficients relative to Wind Production (Daily) Vs Meteorological Using Pearson’s Coefficient

Correlation Coefficients relative to Production Vs Metrological Using Pearson’s Coefficient			
Tested Variable	Metrological reference	Metrological reference	Correlation Value in Percentage
Wind (Daily)	D (Correlation Day (Daily))	Avg(WS)	77.12%
		Min (WS)	74.96%
		Max (WS)	63.69%
		Avg (WD)	40.28%
		Min (WD)	36.07%
		Max (WD)	15.42%
	D-1 (Correlation Day-1 (Daily))	Avg(WS)	42.98%
		Min (WS)	44.90%
		Max (WS)	39.46%
		Avg (WD)	28.03%
		Min (WD)	28.65%
		Max (WD)	12.97%
	D-2 (Correlation Day-2 (Daily))	Avg(WS)	35.11%
		Max (WS)	29.16%
		Avg (WD)	25.65%
	D-3 (Correlation Day-3 (Daily))	Avg(WS)	29.45%
		Max (WS)	23.49%
		Avg (WD)	23.67%
	D-4 (Correlation Day-4 (Daily))	Avg(WS)	27.11%
		Max (WS)	20.05%
		Avg (WD)	20.87%
	D-5 (Correlation Day-5 (Daily))	Avg(WS)	25.78%
		Max (WS)	16.88%
		Avg (WD)	17.17%
D-6 (Correlation Day-6 (Daily))	Avg(WS)	23.21%	
	Max (WS)	13.21%	
	Avg (WD)	16.25%	

Initially, from the above section will be briefly confirmable, the entries with the reference “D-1” would be used as inputs for the model 1. However, after checking that correlation of data only higher correlation were consider for each type of variable Avg WS (D), Min WS (D), Max of WS (D), Avg WS (D-1), Min WS (D-1)For every Input Models there were eliminated or chopped off variables. This eliminated will take place until the least correlation variable of a class type exist

6.5 Section Summary

In this entirety, chapter detail clarifies approximately the information collection and tried with the grouped variable to bring almost grouped set of data that leads advances to achieve the objective with more exactness. Arranged of factors of information were organized for each innovation the information factors were distinctive for each show that has been created. The arranged variable which was utilized in each show would be concisely seen in up and coming chapter.

Assorted factors of authentic and Meteorological were analyzed and examined concisely, and afterwards, this analyzing of information would offer assistance to get it the behavior grouped information variable. This information variable was having an exceptional impact on the improvement of models for each innovation. This examination was performed with Pearson's coefficient, at that point assessing the factors, as well as their authentic references.

After assessment of authentic and meteorological information variable by Pearson coefficient as it were information variable of tall relationship the particular, that sort information variable was well organized for input/ test information gathering. This concise information gathering it will be seen in up and coming section

7. Neural Network Design and Testing

ANNs were situated on mathematical models, have been applied to production forecasting in a liberalized market environment. This strategy is in fact, the foremost predominant production forecasting method within the writing. This method works by learning calculation (Back-Propagation and Levenberg-Marquardt the two most prevalent algorithms) [48]

According to Kenji Suzuki brief explanation stating that the primary process to design network basically undergone two major steps first is to train the network with input and target variables. The second major step is to simulate the network. In this section a detail study to design network. To attain the proposed Model output ANNs were used for predictions with the usage of MATLAB® software. In the beginning, an initial test was accomplished, to choose the tool of this software that it would be used to train and stimulate by creating neural networks. This test was accomplished with the motive, trying to reduce the errors of the models. This method of creating, train, and simulate can be accomplished by two ways which are shortlisted below.

- I. The first one used the nftool (neural fitting tool) for the creation of the networks with which the model would be tested and then with in the same network existing is created in another tool of the same software, the nntool (neural network tool) for the simulation results
- II. In the second method, both the creation of the networks and their simulation would be done in nntool. For each of the methods, 5 output was generated with 3-10 times drills before the simulation. After which the MAPE of each simulation is calculated like the average of those outputs[49].

Indeed, although both test networks have comprised comparable to each other to those utilized for the model in practice would not be attainable, for forecast requires numerous numbers of variables. Although there was a minor difference, the method with a little error, being selected was method second.[48]. In the window neural network data manager nntool input data variable, sample data variable, and target data variable utilized as training data set. In order forward to follow organize, to begin with, select the import from the window which leads to a new window which is shown below

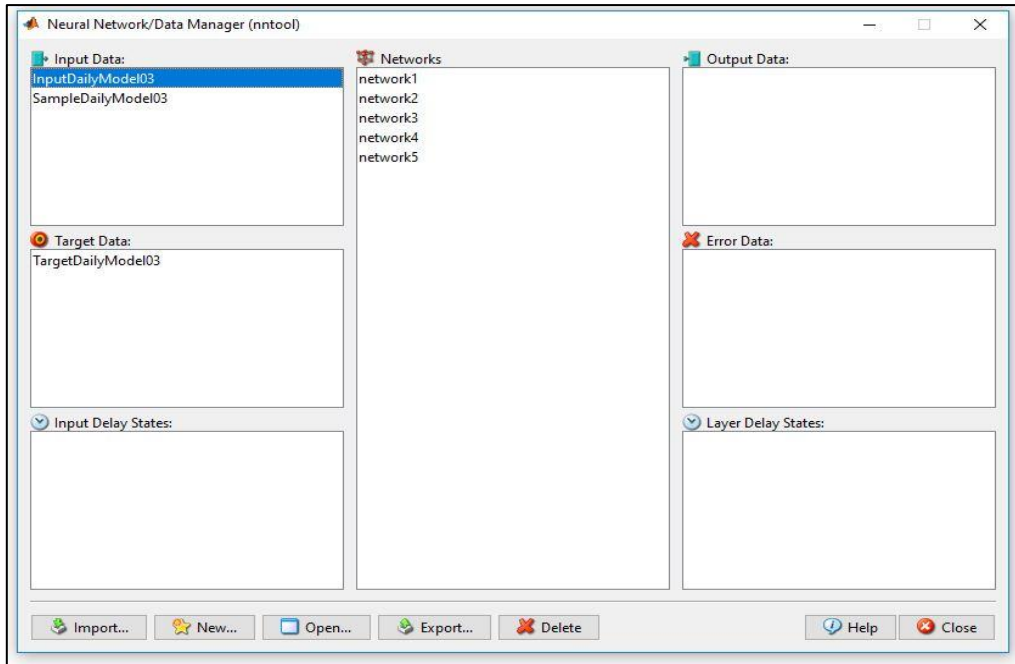


Figure 18 Neural Network data manger nntool

All simulation and networks generated were created in this dissertation according to the second method, with the algorithm feedforward backpropagation, in which the information flow occurs in a unidirectional way, advancing from the input neurons to the output neurons

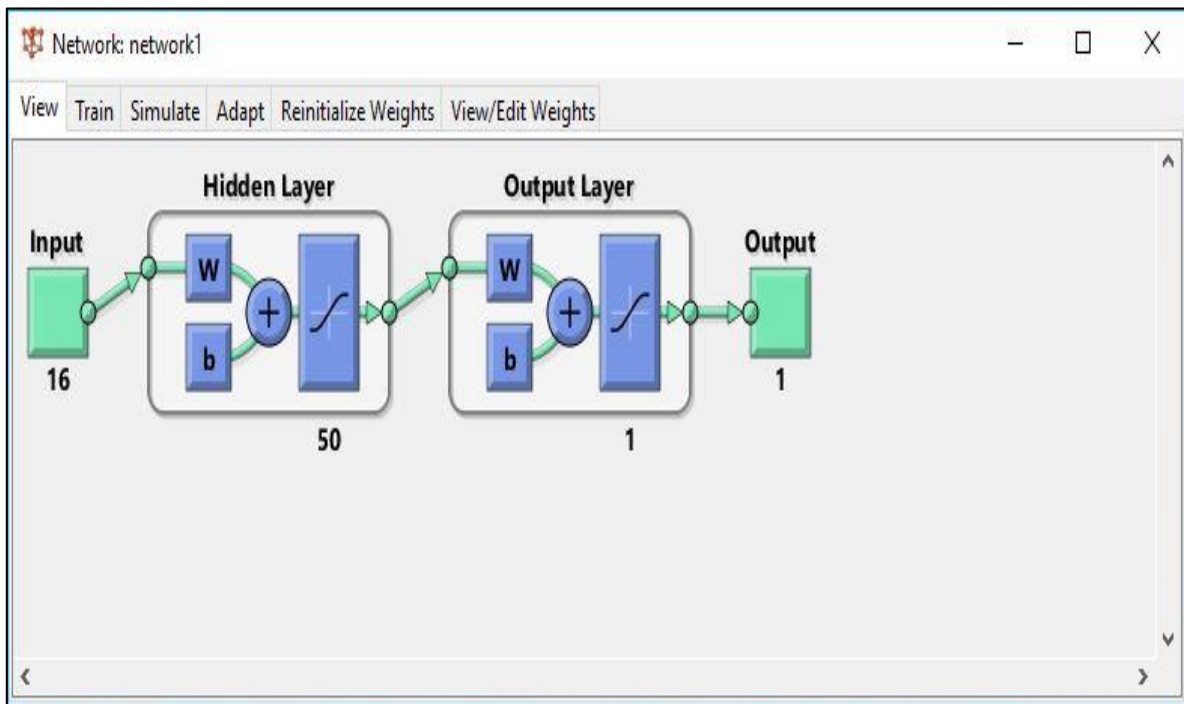


Figure 19 -Example of the structure of a neural network in MATLAB's nntool tool

According to the architecture selected for all networks developed in this dissertation, they are networks multilayer, or multi-layer perceptron's and in this type of networks, there is a layer known as the hidden layer containing a variable number of neurons this being a fact that makes this kind of networks very used. Two-layer nets were used, there is no universal

rule for selecting the number of neurons in the layer the rule that was used defines the use of $(3*n+2)$, neurons where n is the number hidden layer [10]. An example of the application of this rule to a network created with 16 entries, being that applying the mentioned rule the number of neurons in the hidden layer is 50.

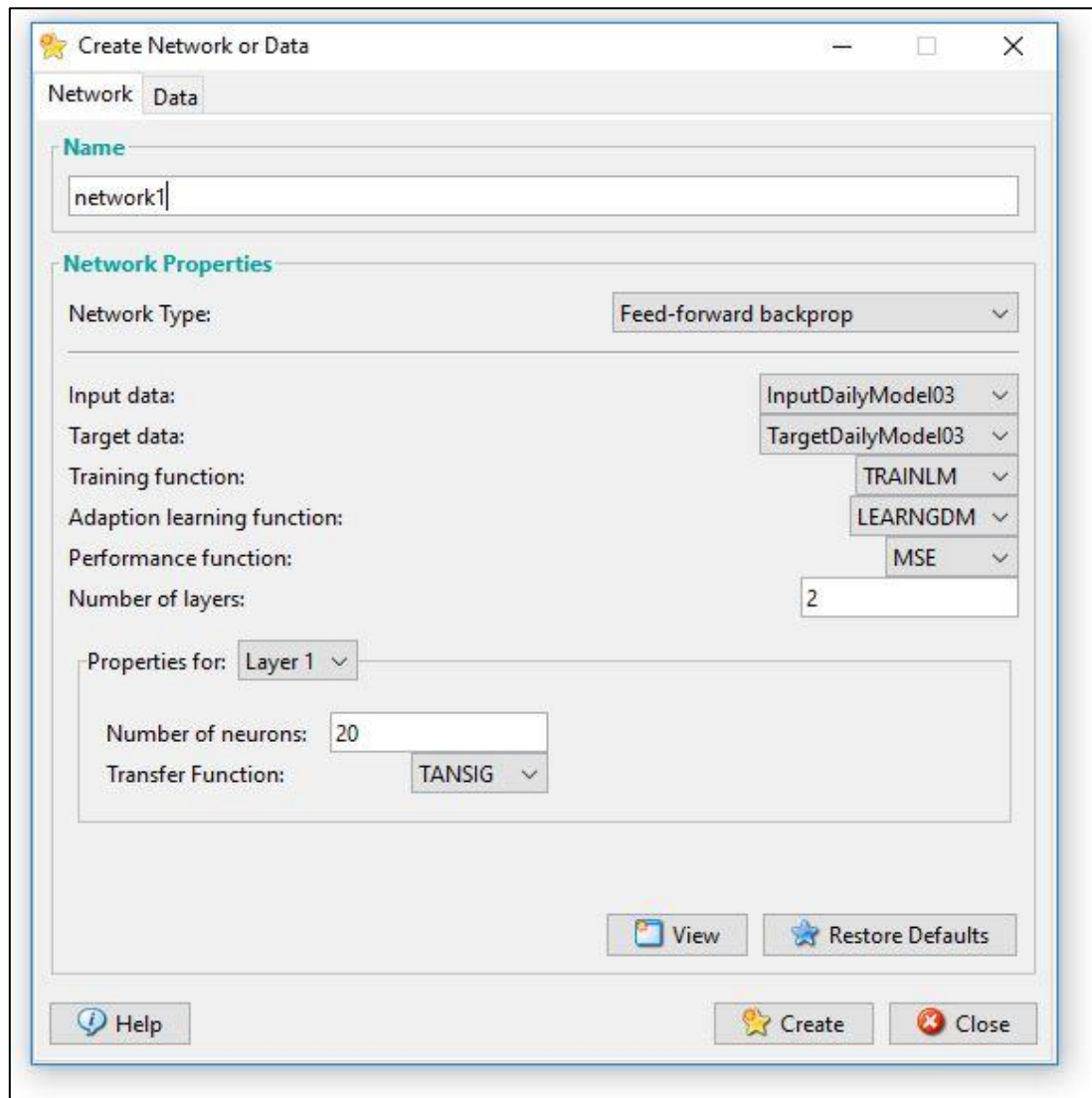


Figure 20 Creation of networks parameter

In the rest of parameters to be defined in the networks, Levenberg-Marquardt backpropagation, and as functions of adaptation of learning and performance a learned (gradient descent with momentum weight and bias learning function) and MSE (mean-squared error), respectively. In layer 1 (hidden layer) the TANSIG function was used as a transfer function and in layer 2 (output layer) the Purelin function[49, 50].is represented in above

To use the evaluation of the networks to be built the data collected and processed for the construction were divided. In the hourly and daily models, the data used were analogous to the time period which is shortlist in Table 4. These data were distributed so that their division was about 80% for training data, and 20% for test data (sample data). By this process, the data training would serve to the neural network train that one may to “teach it” for the given data presented.

This training process would then be simulated, with corresponding data of 20% representatives of the test data (Sample data). The sample data for tested can be selected in a random choice of 10 weeks but in this dissertation, every 2nd week is selected as a sample or test data for about 10 months (from January to October is about 20% approx. of the total sample from 365 days this is applicable for hourly and daily all data models and technology). In fact, there is a special case for Hydro technology in that sample data were selected from January to August is about 20% approx. of the total sample from 325days, there is phase-out of data in this technology which were briefly explained in the previous section. The 10 weeks of the samples set chosen is shortlisted below.

Table 18 Sample data for selected weeks each Technology

Technology	Year	Date	Month (Every Month 2nd Week)									
			Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct
Hydro	2019	Date	7 to 12	3 to 9	3 to 9	7 to 13	5 to 11	2 to 8	7 to 13	4 to 10	Phase Out Data	
Solar	2018		7 to 13	4 to 10	4 to 10	8 to 14	6 to 12	3 to 9	8 to 14	5 to 11	2 to 8	7 to 13
Wind	2018		7 to 13	4 to 10	4 to 10	8 to 14	6 to 12	3 to 9	8 to 14	5 to 11	2 to 8	7 to 13

These designated training and test sets were used in the hourly and daily models. In this way outline of the methodology developed to choose the best model for each of the time-period were considered which is elaborate shown in the flow chat beneath.

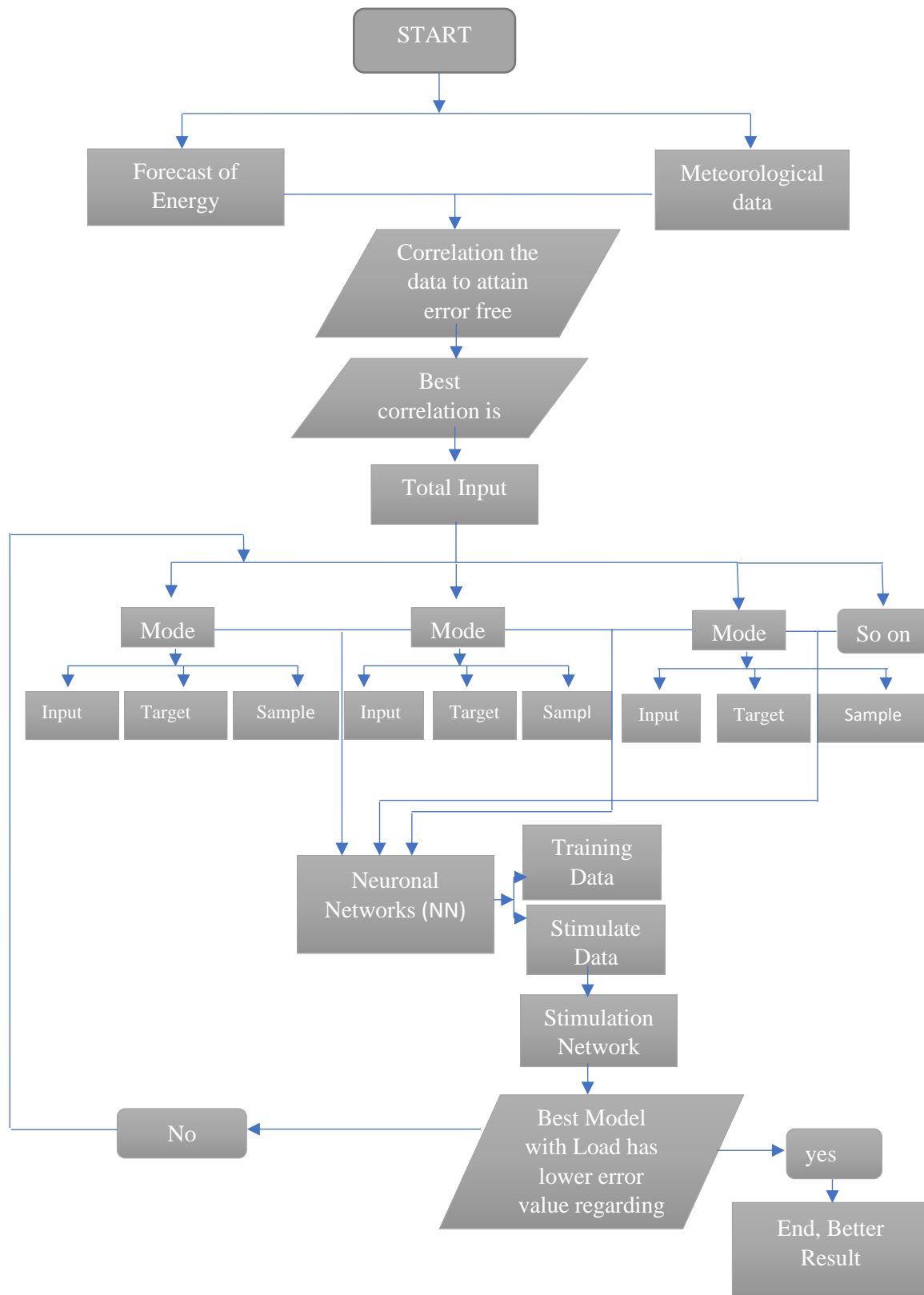


Figure 21 Methodology developed to choose the best model

To implement all the case studies of this dissertation, a Laptop was used with an Intel (R) Core (TM) i7-6500U CPU @ 2.50 GHz 2.59 GHz, 12 GB RAM and with a Windows 10 Home 64-Bit operating system.,x64-based processor.

7.1 Forecasting Model Evolution

For each model, there is more recent information available which means more variable to be chosen from the correlation of variables were undertaken in the based on previous plots section above. This also makes the error tolerance being lower for every model and the same format is germane to all the technology. This development was later compared between hourly and daily which every model has less tolerance error leads forward to attain the best or outstanding model.

7.1.1 Forecasting Model Training Error or Overfitting

As Mentioned before section during the drills 3-10 times while creating the network. There met to be the hazard of error in regression line to attain saturation point the weights of the training network is reinitialized and the network is again trained drills 3-10 times. This mechanism is replicated until regression line attains the saturation line. While accomplishing the objective of this dissertation there were few cases of regression line error which is shown beneath

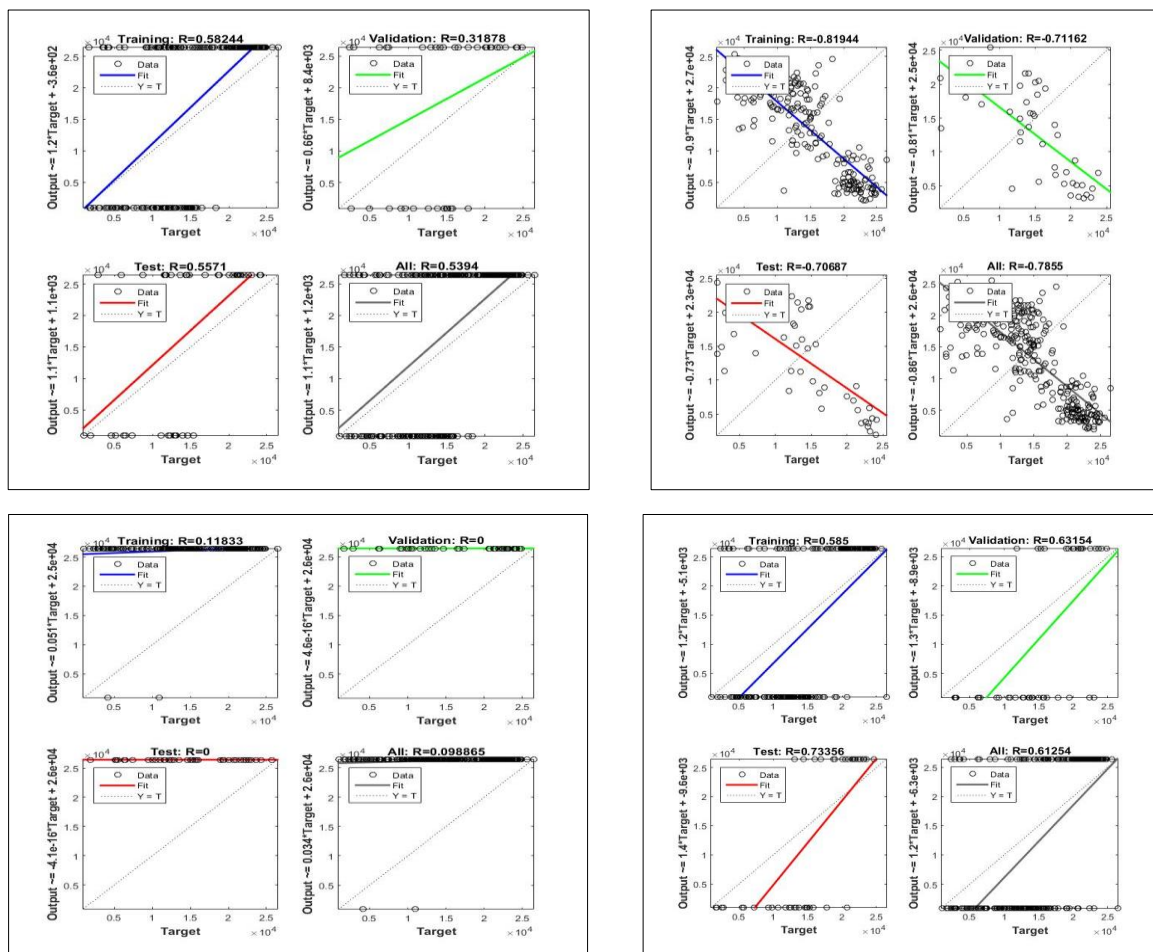


Figure 22 Example for Regression line error

7.1.2 Model Parameters

To accomplish forecast models all the care to be taken into consideration, which leads to a guarantee in making daily and hourly, in addition, tests were carried out for all the models. Before proceeding to the training and simulation of the network of the different models together with each technology, tests with standard parameters to be used and the number of training to be performed within the simulation of all the models[49]. For these tests, the Matlab ® tool is utilized. These aspects were assessed concurring to the ensuing investigation of the result of the MAPE errors. The all the tests that were carried out for the models of all technology the parameter utilized within the training of network is shown within the underneath figure.

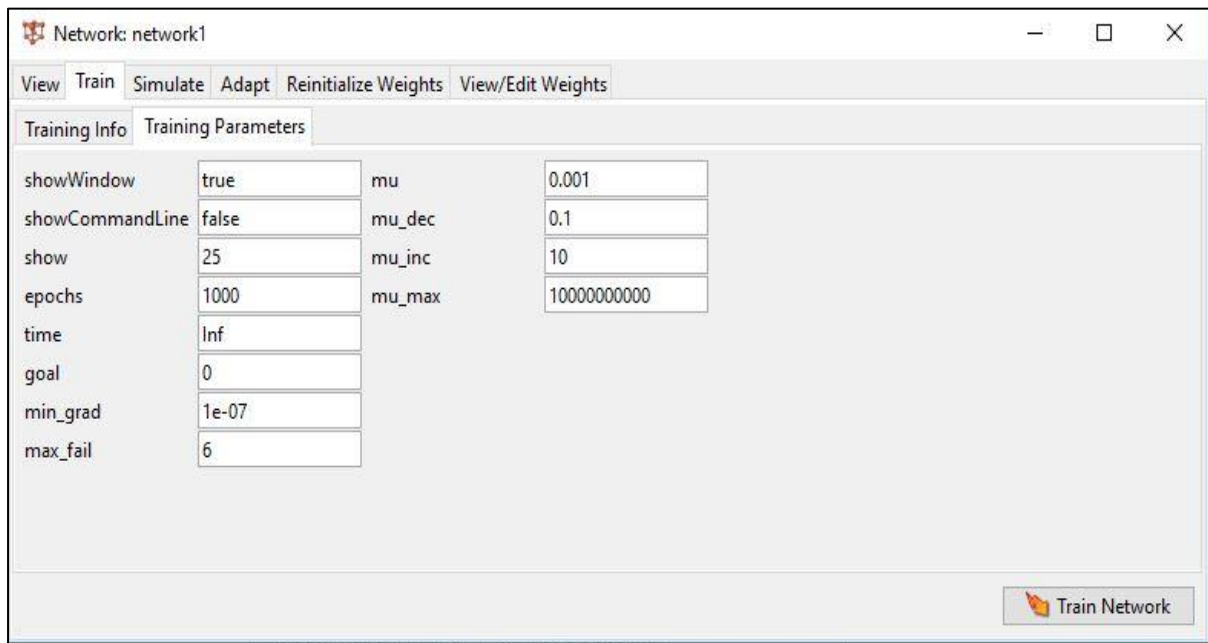


Figure 23 Training Parameters

This parameter utilized within the training of the network to Maximize the accuracy of the models to be created. this impact, five models were designed with the same parameter values. This may offer assistance the network to achieve solidness and the greatest number of validation checks take put each time network is trained with 3-10 times of drills without execution change (Max_fail) and minimum magnitude slope (Min_grad)[49].

These values are utilized finalize to the training of the neural networks. Each model was tested three -ten times to have more than one output to test their result more effectively. Five systems were developed for each model to form the specified number of outputs having been executed five workouts in each one before the subsequent. Simulation of their results, which is deliberately shown within the table below

Table 19 Results of the parameters models created

Technology	Network Parameters	MAPE	Min_grad	Max_fail	epochs	mu	mu_dec	mu_inc	mu_max							
Hydro(hourly)	Model 01	48.47%	1.00E-07	6	1000	0.001	0.1	10	1E+10							
	Model 02	38.84%														
	Model 03	40.41%														
Hydro(Daily)	Model 01	24.78%														
	Model 02	23.59%														
	Model 03	21.44%														
Solar(hourly)	Model 01	44.91%								1.00E-07	6	1000	0.001	0.1	10	1E+10
	Model 02	38.00%														
	Model 03	38.50%														
	Model 04	49.97%														
Solar(Daily)	Model 01	31.71%														
	Model 02	25.24%														
	Model 03	26.80%														
	Model 04	23.72%														
Wind(hourly)	Model 01	480.29%	1.00E-07	6	1000	0.001	0.1	10	1E+10							
	Model 02	984.40%														
	Model 03	1161.05%														
Wind(Daily)	Model 01	63.63%														
	Model 02	65.74%														
	Model 03	55.31%														
	Model 04	58.37%														
	Model 05	75.18%														

As shown in Table 19 the networks that proved to be more accurate in this test were those that were trained only with fixed parameters chosen and tested for all the technology. The test described has been accomplished to minimize the error obtained in the simulation of the forecast the models to be developed for this purpose and the subsequent analysis of the MAPE error for each of these. For the development of these models, the variables selected were used based on the studies carried out in section 6.

Table 20 Test results regarding the best model based on each technology and respective five networks of training used

Technology	Best Model Results	Avg.MAPE	Network Training	MAPE
Hydro(hourly)	Model 02	38.84%	Network 1	47.22%
			Network 2	46.61%
			Network 3	46.70%
			Network 4	39.70%
			Network 5	38.61%
Hydro(Daily)	Model 03	21.44%	Network 1	25.68%
			Network 2	21.97%
			Network 3	26.79%
			Network 4	24.79%
			Network 5	26.33%
Solar(hourly)	Model 02	38.00%	Network 1	49.64%
			Network 2	47.12%
			Network 3	46.30%
			Network 4	46.82%
			Network 5	50.44%
Solar(Daily)	Model 04	23.72%	Network 1	26.54%
			Network 2	24.68%
			Network 3	24.37%
			Network 4	26.18%
			Network 5	25.95%
Wind(hourly)	Model 01	480.29%	Network 1	363.87%
			Network 2	263.92%
			Network 3	300.07%
			Network 4	519.32%
			Network 5	1083.38%
Wind(Daily)	Model 03	55.31%	Network 1	56.67%
			Network 2	62.52%
			Network 3	51.58%
			Network 4	56.76%
			Network 5	55.36%

Of the models created and tested which obtained a smaller MAPE can be seen in Table 19 consequently chosen the same parameter used for remaining all simulation in the hourly and daily production forecasting models. The next test accomplished was the number of training sessions to be performed in each network before proceeding the simulation of that particular network.

In this dissertation number of times, the network is trained if 3,5 and 10 times within that when validation check of the regression line is high the training stops and the network is simulated or the network is trained 3,5 and 10 times and later it simulates either one way it is executed. This method of training and simulation was carried out for all the five networks for each model. This was followed in all other technology too. Subsequently, for each model test, five output was generated the average value were consider for that model output. The best model results networks relative to the MAPE is shown in Table 20. Network Training regression graphical representation is shown below.

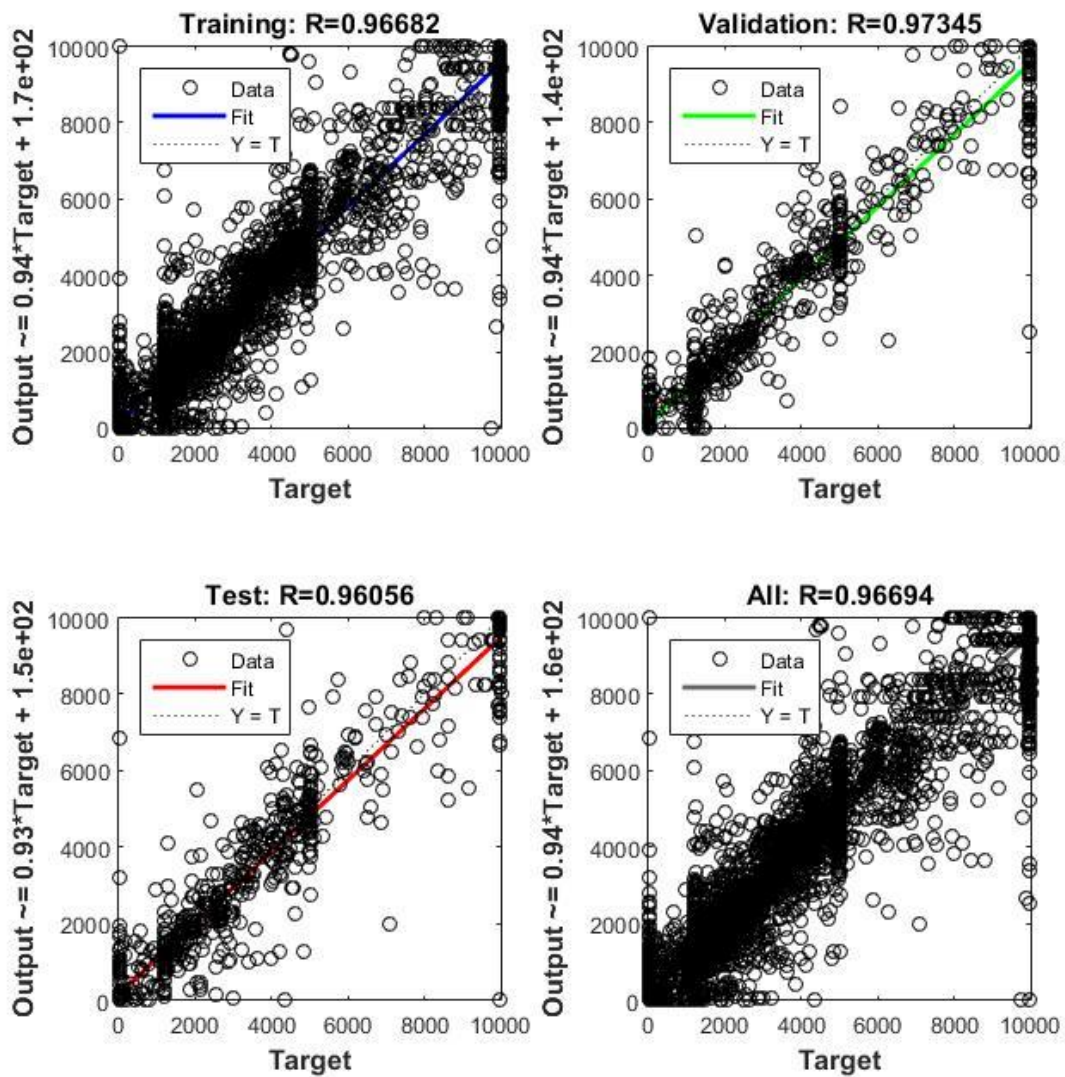


Figure 24 Example of a regression graphical representation, from the nntool

To the inclusion of the selected variables in respective models essentially based on the correlation values which were calculated by the Pearson coefficient, as referenced in the previous section also including the input identifying variables for each model.

8. Technology's and Respective Results

8.1 Hydro Technology Input variables

Before all the variable input was considered this technology is a special scenario. In this hydro technology, there was very little time where the values have eliminated the values due to no production in that period of time. This no production duration can be due to lack of rainfall in that time of period, or overwhelming maintenance within the same period that could moreover attainable. Comprehension to eliminate this variable might lead to critical towards accomplishing the accuracy of the result. In this dissertation, eliminated factors period for 6/9/2019 to 15/10/2019 this can be as it were applicable on for this technology only.

8.1 1 Hydro Technology Input Variables

To begin with going in spite of the fact that the hydro technology hourly production model 01 (MPH 01) input variables entries considered most momentous within the enlargement of the historical energy production (D-1),(D-2) and (D-3) together with that Meteorological data variable such as P(D-1), Sum P(D-5), Min P(D-1), Avg H(D-5) and Min H(D-1) with a Pearson coefficient value of around 0.85,0.78,0.69 and Meteorological information Pearson coefficient 0.45,0.56,0.45,0.68 and 0.68 individually. There is one more variable supplemental is month variable indeed in spite of the fact that month data variable correlation is low it stabilizes and benefits the network. This all variable was partitioned 80:20 proportion because it briefly clarified within the above area in this 80% is input and 20% is the samples to be tested. This classification is regarded for each and each model besides their respective technology.

One of the foremost critical things is that whereas considering the variables of past historical and past Meteorological data most behind variable with high correlation data input is considered from that and underneath the data would be taken as thought over esteem would be eliminated. For illustration, MPD 01 comprises the diverse high correlation of the variable was considered as input in that slightest historical or most past Meteorological data, past day-6 esteem sum P(D-6) and Avg (D-6) in spite of the fact that P(D-1), Min P(D-1), Min H(D-1) are higher still the esteem is selected from 7 days and underneath in this way the data variable was chosen is clarified with Figure 25.

Date	Month	Actual Production (A)	Previous Historical Value						Metrological Data				
			D-1	D-2	D-3	D-4	D-5	D-6	P (D-1)	Min P (D-1)	Min H (D-1)	Sum P (D-6)	Avg H (D-6)
01-01-2019	1	X											
02-01-2019	1	X	X						X	X	X		
03-01-2019	1	X	X	X					X	X	X		
04-01-2019	1	X	X	X	X				X	X	X		
05-01-2019	1	X	X	X	X	X			X	X	X		
06-01-2019	1	X	X	X	X	X	X		X	X	X		
07-01-2019	1	X	X	X	X	X	X	X	X	X	X	X	X
08-01-2019	1	X	X	X	X	X	X	X	X	X	X	X	X
09-01-2019	1	X	X	X	X	X	X	X	X	X	X	X	X
10-01-2019	1	X	X	X	X	X	X	X	X	X	X	X	X
11-01-2019	1	X	X	X	X	X	X	X	X	X	X	X	X

Figure 25 Example selecting of data variables

Within the second model of the Hydro Technology model, few of the variables were eliminated by holding the high correlation esteem of each sort variable. Here each sort variable implies for illustration MPH 01 D is historical esteem, P pluviosity M-month, and H- humidity were all alluded to as a sort variable.

Table 21. Hydro technology hourly production forecasting models (MPH) entries variable with respective to correlation value in percentage

Types of Input Variable and respective Correlation Percentage		
Models (hourly)	Historical Production reference	Correlation Value in Percentage
Input Model 01 (hourly)	D-1	85.49%
	D-2	78.34%
	D-3	69.49%
	Metrological reference	
	P(D-1)	66.48%
	Sum P(D-5)	56.08%
	Min P(D-1)	45.37%
	Avg H (D-5)	68.39%
	Min H(D-1)	69.60%
	Month (M)	23.28%
Input Model 02 (hourly)	Historical Production reference	
	D-1	85.49%
	D-2	78.34%
	Metrological reference	
	Sum P(D-5)	56.08%
	Min P(D-1)	45.37%
	Avg H (D-5)	68.39%
	Min H(D-1)	69.60%
Month (M)	23.28%	
Input Model 03 (hourly)	Historical Production reference	
	D-1	85.49%
	Metrological reference	
	Sum P(D-5)	56.08%
	Min P(D-1)	45.37%
	Avg H (D-5)	68.39%
	Min H(D-1)	69.60%
Month (M)	23.28%	

In this third model two-variable from the diverse classes were eliminated in this way the network steadiness and decrease the total error rate. The variable information that was taken into consideration is shortlisted in

Within the second model of the Hydro Technology model, few of the variables were eliminated by holding the high correlation esteem of each sort variable. Here each sort variable implies for illustration MPH 01 D is historical esteem, P pluviosity M-month, and H- humidity were all alluded to as a sort variable.

Table 21. In the third show, one variable is eliminated from the past model and the remaining data variable is considered. The variable which was considered is shortlisted in

Coming to the daily models it is pretty much comparative to hourly models of hydro technology input variables entries considered most momentous within the extension of the historical energy production (D-1),(D-2) and (D-3) together with that Meteorological data variable such as P(D-1) Sum P(D-6) Min P(D-1) Avg H (D-6) Min H(D-1) and Month (M)with a Pearson coefficient esteem of approximately 0.91,0.82,0.73 and Meteorological data Pearson coefficient 0.49,0.60,0.49,0.73,0.75 and 0.25 respectively.

In the second model of hydro technology model, few of the variables were eliminated or eliminated by retaining the high correlation value of each type variable. Here each type variable means for example MPH 01 D is historical value, P pluviosity M-month and H- humidity were all referred to as a type variable. In this model two-variable from the different class were eliminated in this way the network stability and reduce the total error percentage. The variable data were taken into consideration is shortlisted in Table 22.

Table 22. Hydro technology Daily production forecasting models (MPD) entries variable with respective to correlation value in percentage

Types of Input Variable and respective Correlation Percentage			
Models (Daily)	Historical Production reference	Correlation Value in Percentage	
Input Model 01 (Daily)	D-1	91.95%	
	D-2	82.95%	
	D-3	73.73%	
	Metrological reference		
	P(D-1)	49.28%	
	Sum P(D-6)	60.53%	
	Min P(D-1)	49.10%	
	Avg H (D-6)	73.94%	
	Min H(D-1)	75.32%	
	Month (M)	25.26%	
Input Model 02 (Daily)	Historical Production reference		
	D-1	91.95%	
	D-2	82.95%	
	Metrological reference		
	Sum P(D-6)	60.53%	
	Min P(D-1)	49.10%	
	Avg H (D-6)	73.94%	
	Min H(D-1)	75.32%	
Month (M)	25.26%		
Input Model 03 (Daily)	Historical Production reference		
	D-1	91.95%	
	Metrological reference		
	Sum P(D-6)	60.53%	
	Min P(D-1)	49.10%	
	Avg H (D-6)	73.94%	
	Min H(D-1)	75.32%	
Month (M)	25.26%		

In the third model, one variable is eliminated from the previous model and the remaining data variable is considered. The variable which was considered is shortlisted in Table 22

8.1.2 Hydro Technology Result and Discussion

After all the input variable were consider based on the above specification then the network is simulated respective output is generated. This output variable used as an input variable to calculate MAPE. The model that obtained a smaller MAPE in this hydro technology was MPH 02 with 38.84% in the hourly model and in MPD 03 with 21.44% in the daily model which is shortlisted in the below table

Table 23 Hydro Technology final model dropped values from

TECHNOLOGY	MODEL	MAPE
Hydro(hourly)	Model 01	48.47%
	Model 02	38.84%
	Model 03	40.41%
Hydro(Daily)	Model 01	24.78%
	Model 02	23.59%
	Model 03	21.44%

The mean value obtained for each input hour in its stimulation were used to test the model to verify its errors. The model which obtained lower error was considered as the best model for that particular technology. Each simulation was carried out considering the criteria of previously determined with 5 output. Later average output error obtained is simulation resulted from MAPE. As expected the error value is higher than the previous model this justified since the load variable to forecast itself has associated with the error. However, the value remains below the MAPE of the best model making these the models selected for production forecast in the daily market. Finally, it was used explanation to proceed with management of the error.

Applying this process it was possible to obtain a reduction in the error of the final model which is shown in Table 23. Summary of the method carried out to forecast the production of electricity in the development based on hourly and daily models

8.1.3 Hydro Technology Result for Hourly Model

The correlation of the variable so obtained was not auspicious result as expected. This correlation of the variable was carried out by using the Pearson coefficient. Comparing with other models os hourly this obtained value of model no 02 is better than other models of hourly. Then the obtained results of the forecast value were related to the actual production value along with test set variables. This information is graphically represented in Figure 26.

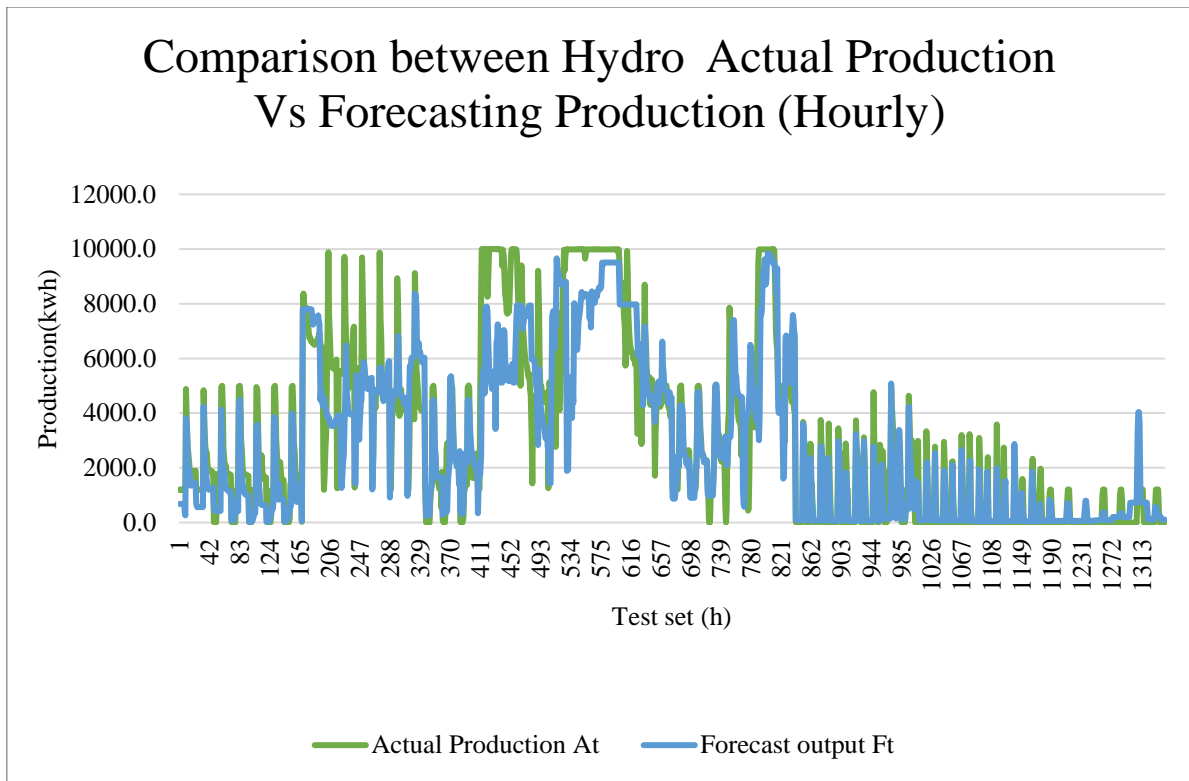


Figure 26 Comparison between the results obtained and actual production values (hourly model)

The Figure 26 symbolize information about hydro technology production in Portugal over a time period a year (2019) summarise the information by selecting and reporting the main features to make comparisons where relevant. Figure 26 for actual production, forecast production was almost similar in the certain time period and move far away in a certain time period at around 38.8 per cent of error.

The actual production and forecast output possibilities frontier is constructed by plotting all the possible combination of the output represented in Figure 26. From the information represented in the graph it is clear that the test set point at 521 to 581 the two lines is a little bit distant this has occurred when error distinguishes between actual and output production is higher. When it comes to testing set point 1 to 161 the line of the graph are close enough to one other this represents the forecasting variable at the point good with the smaller error were found.

The hourly model it seems to have a greater correlation in following the Actual production when there are peaks once again the week of January 6 to 12 was isolated 2019 with the results obtained in the simulation of this forecast model with a view to comparison with the values confirmed that week for production Figure 27.

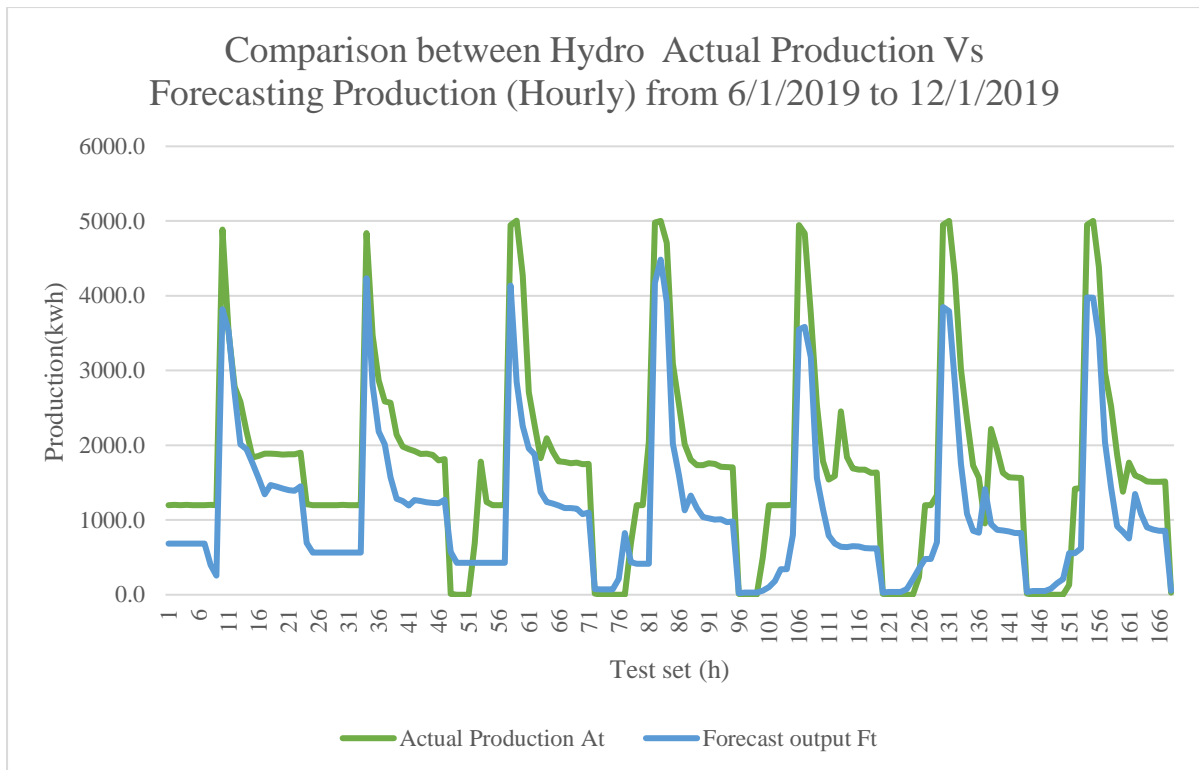


Figure 27 Week from January 6 to 12, 2019, with the comparison of the Actual values of production and obtained outputs (Hourly model)

Figure 27 information about hydro technology production in Portugal over the time period a year (2019) summarises the information by selecting and reporting the main features to make comparisons where relevant.

Here you can see the difference that exists between the outputs that are generated by the hourly model developed and the value analyzed along with actual energy production. As you can see the values generated by the model, usually slightly below the actual production values, with respect to the week that was chosen.

8.1.4 Hydro Technology Error Obtained (Hourly Model)

The histogram graph represents the error difference between the actual production and simulated 5 output avg forecasting production of the hourly model. This graph gives information about colossal of average error confirming with absolute production value. when the deviation of value between the average forecast value and actual production is high this leads to error occurrence this analysis of the data is shown in the below Figure 28, Figure 29. Represented the histogram of the relative obtained output production errors can be established.

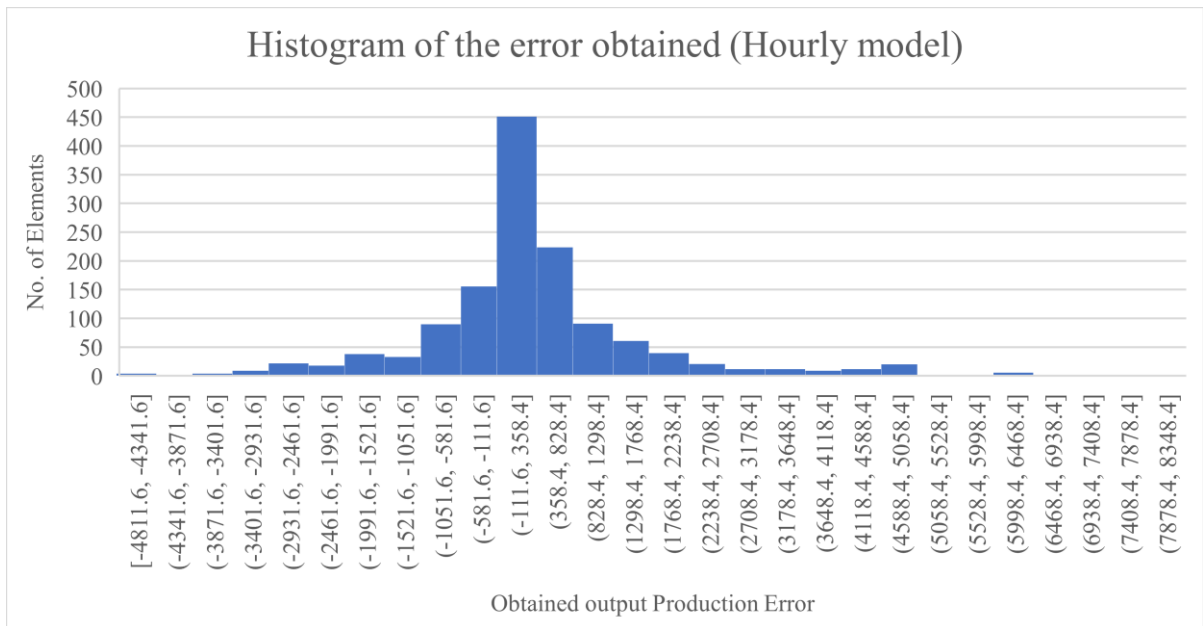


Figure 28 Example Histogram of the Production error obtained (Hourly model)

The graph shows information about production errors obtained will creating and executing the model centred on the interval containing zero and are relatively distributed around these values. It represents relative error over the sample of the test set used in Figure 29.

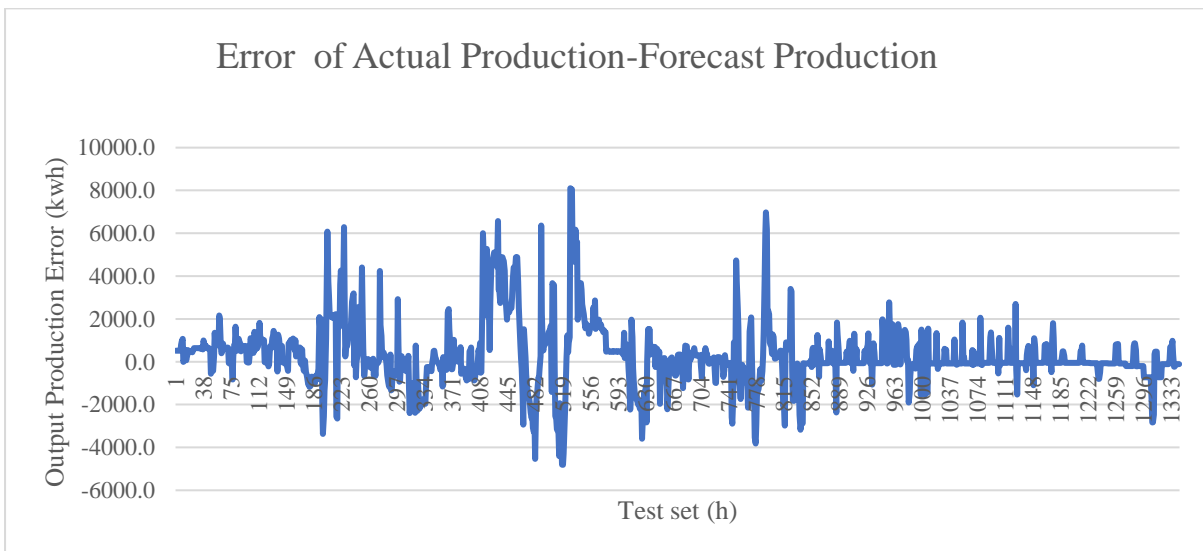


Figure 29 Example Obtained output production error and standard deviation (Hourly model)

The Figure 29 graphical representation gives information about the highest and lowest error obtained from output production. These variables are compared with the hourly test set variable. the error of actual – output production is the peaks that occur are represented peaks in the graph. The colossal error found was around -4811.6 kWh and may be due to the great flaw pointed to the method used for the performance calculation of the MAPE. This collapse is due to obtaining Actual- forecast output production values of zero or close to zero which make the error found to be very high regardless of the error absolute value. The results obtained and calculations performed in this model can be seen in the Annex1.

8.1.5 Hydro Technology Results for Daily Model

The correlation of the variable so obtained was not auspicious result as expected. This correlation of the variable was carried out by using the Pearson coefficient. Comparing with other models os hourly this obtained value of model no 03 is better than other models of hourly. Then the obtained results of the forecast value were related to the actual production value along with test set variables. This information is graphically represented in Figure 30. However, the values are able to follow their evolution despite its volatility. This can be seen in the daily model the difference that exists between the outputs that are generated in the hourly model developed is comparable small then daily model. Daily model representation graph shown beneath with the average MAPE verified for this model in the test set.

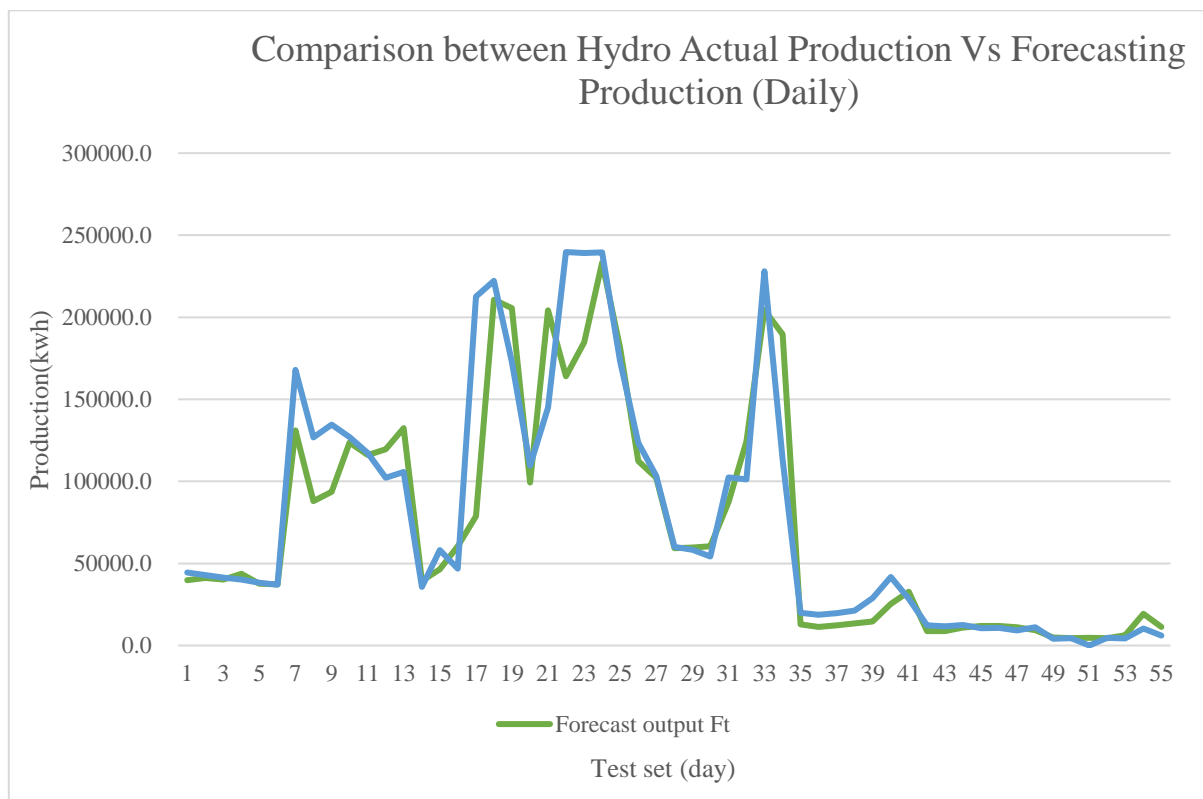


Figure 30 Comparison between the results obtained and actual production values (Daily model)

The Figure 30 symbolize information about hydro technology production in Portugal overtime period a year (2019) summarise the information by selecting and reporting the main features to make comparisons where relevant. Figure 30. for actual production, forecast production were almost similar in a certain time period and move far away in a certain time period at around 21.44 per cent of error.

The actual production and forecast output possibilities frontier is constructed by plotting all the possible combination of the output represented in Figure 30. From the information represented in the graph it is clear that the test set point at 23 to 25 the two-line is a little bit distant this has occurred when error distinguishes between actual and output production is higher. When it comes to testing set point 1 to 7 the line of the graph are close enough to one other this represents the forecasting variable at the point good with the smaller error were found.

The Daily model it seems to have a greater correlation in following the actual production when there are peaks once again the week of January 6 to 12 was isolated 2019 with the results obtained in the simulation of this forecast model with a view to comparison with the values verified that week for electricity production Figure 31.

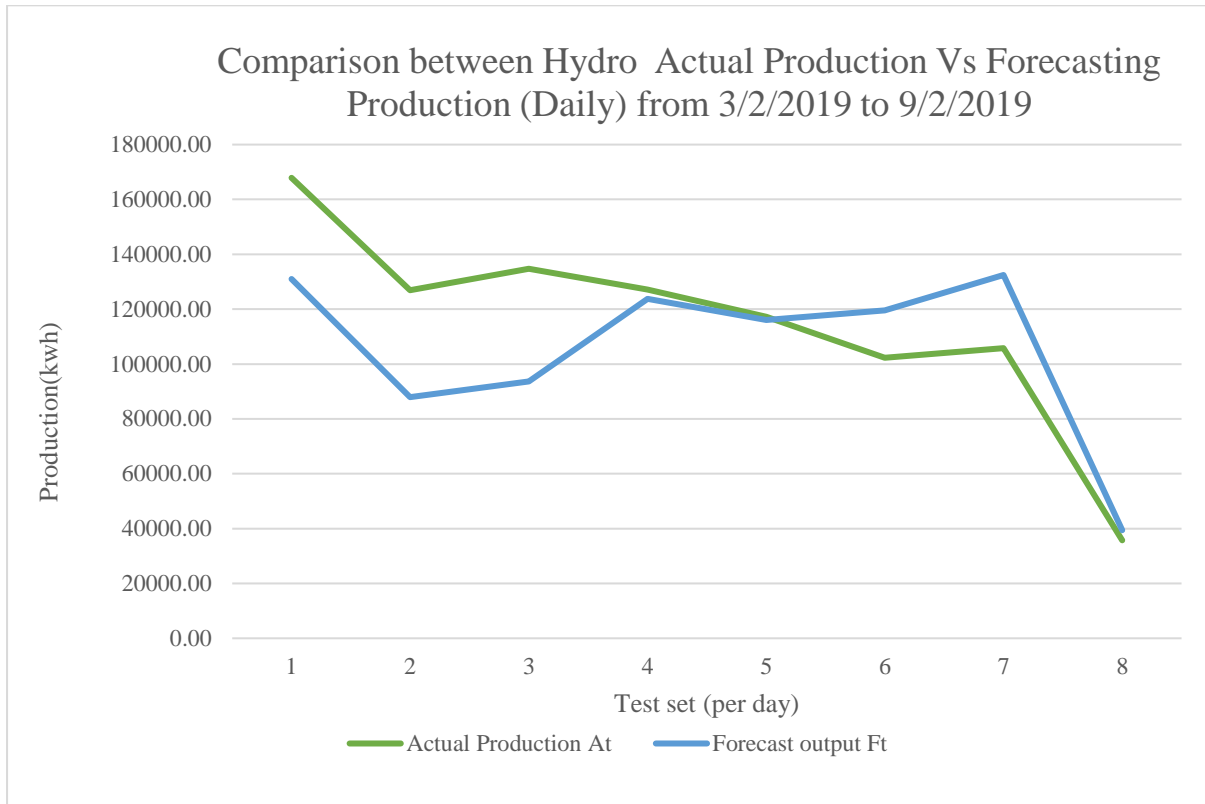


Figure 31 Week from February 3 to 9 (2019), with the comparison of the Actual values of production and obtained outputs (Daily model)

Figure 31 information about hydro technology production in Portugal overtime period a year (2019) summarise the information by selecting and reporting the main features to make comparisons where relevant.

Here you can see the difference that exists between the outputs that are generated by the daily model developed and the value analyzed along with actual energy production. As you can see the values generated by the model, usually slightly below the actual production values, with respect to the week that was chosen.

8.1.6 Hydro Technology Error Obtained (Daily Model)

The histogram graph represents the error difference between the actual production and simulated 5 output avg forecasting production of the hourly model. This graph gives information about colossal of average error confirming with absolute production value. when the deviation of value between the average forecast value and actual production is high this leads to error occurrence this analysis of the data is shown in the below Figure 32, Figure 33. Represented the histogram of the relative obtained output production errors can be established.

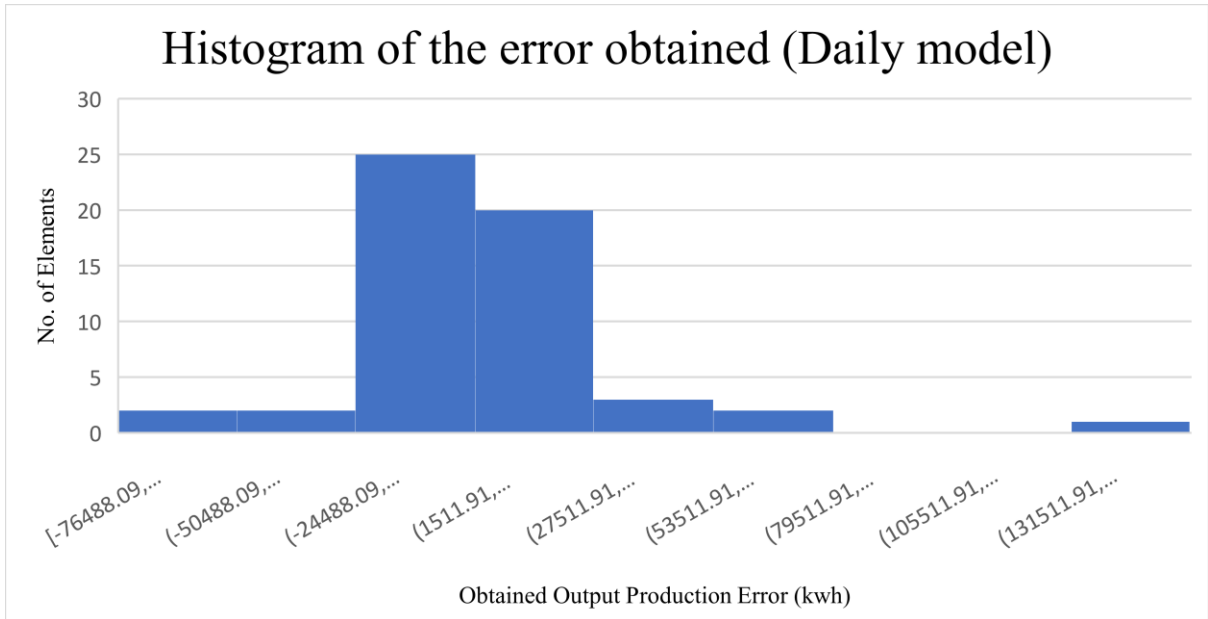


Figure 32 Example Histogram of the Production error obtained (Daily model)

The graph shows information about production errors obtained will creating and executing the model centred on the interval containing zero and are relatively distributed around these values. It represents relative error over the sample of the test set used in Figure 33.

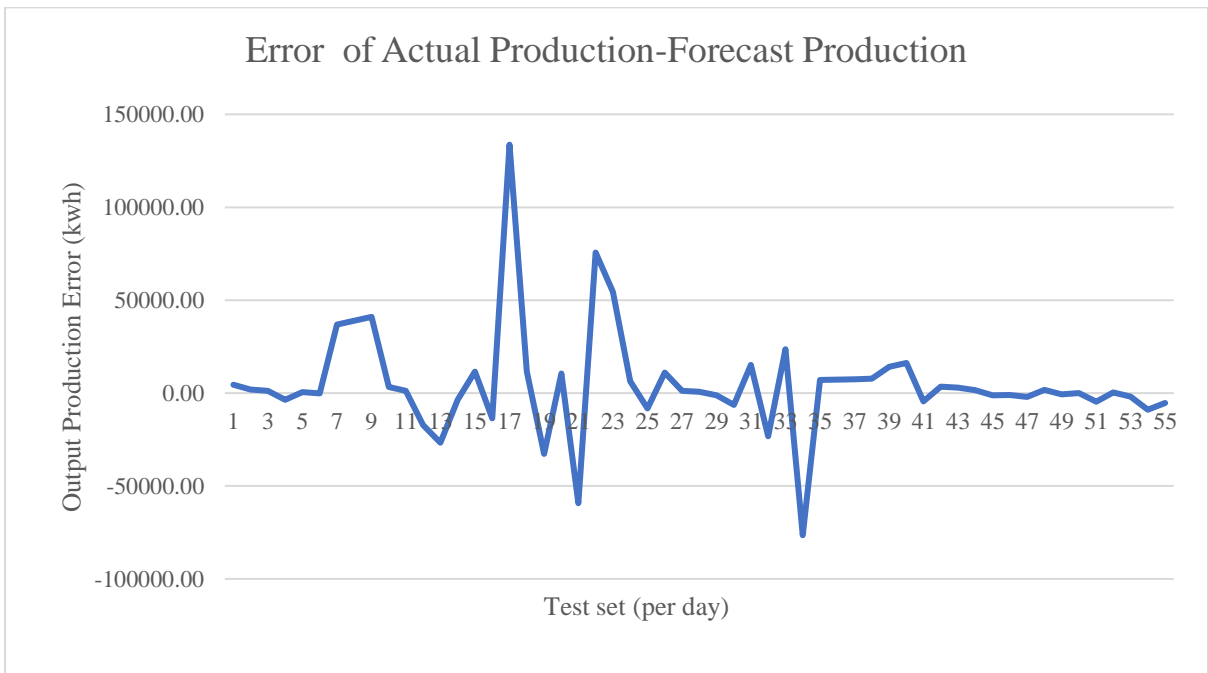


Figure 33 Example Obtained output production relative error and standard deviation (Daily model)

The Figure 33 graphical representation gives information about the highest and lowest error obtained from output production. These variables are compared with the daily test set variable. the error of actual – output production is the peaks that occur are represented peaks in the graph. The colossal error found was around -76488.08 kWh and may be due to the great flaw pointed to the method used for the performance calculation of the MAPE. This collapse is due to obtaining actual- forecast output production values of zero or close to zero which

make the error found to be very high regardless of the error absolute value. The results obtained and calculations performed in this model can be seen in the annex 1.

8.2 Solar Technology

In the case of solar technology, the variable such as Historical data and Meteorological data were gathered from the different platform from which variables with high correlation is considered as input. Later tested multiple times to reduces the false-positivity of data set. In this way, data sets were well organized. This correlation of variables was calculated by person coefficient. Input variables identified for each model differ from one other. This variable was carried out to test multiple times which lead forward to attain the least error forecasting models.

8.2 1Solar Technology Input Variables

For the first model of solar technology hourly the variables identified and considered most momentous in the enlargement of the historical energy production (D-1),(D-2), (D-3) (D-4), (D-5) and (D-6) along with that Meteorological data variable such as Max of T (D), Avg I (D), Max I (D), Min T (D-1),Avg T (D-2), Avg-Hh, Avg-Hi_opt, Avg-Hi_m and Avg-Hb(n)_m with a Pearson coefficient value of about 0.88,0.87,0.86,0.86,0.86 and 0.86 Meteorological data Pearson coefficient 0.22,0.21,0.21,0.18,0.21,0.19,0.19 and 0.19 respectively. There is one more variable supplemental is month variable even though month data (0.06) variable correlation is low it stabilizes and benefits network. This all variable was divided 80:20 ratio as it briefly explained in the above section in this 80% is input and 20% is the samples to be tested. This classification is respected for each and every model along with there respective technology.

One of the most significant things is that while considering the variables of previous historical and pervious Meteorological data most behind variable with high correlation data input is considered from that and beneath data would be taken as consideration above value would be eliminated. For example, MPD 01 consist of the different high correlation of variable were considered as input in that least historical or most previous Meteorological data, previous day-6 value are higher still the value is selected from 7 days and beneath in this way the data variable was selected.

Table 24 Solar technology Hourly production forecasting models (MPH) entries variable with respective to correlation value in percentage

Types of Input Variable and respective Correlation Percentage		
Models (hourly)	Historical Production reference	Correlation Value in Percentage
Input Model 01 (hourly)	D-1	88.83%
	D-2	87.54%
	D-3	86.94%
	D-4	86.82%
	D-5	86.49%
	D-6	86.26%
	Metrological reference	
	Max of T (D)	22.45%
	Avg I (D)	21.52%
	Max I (D)	21.03%
	Min T (D-1)	18.69%
	Avg T (D-2)	21.17%
	Avg-Hh	18.49%
	Avg-Hi_opt	19.23%
	Avg-Hi_m	19.06%
	Avg-Hb(n)_m	19.36%
Month (M)	6.13%	
Input Model 02 (hourly)	D-1	88.83%
	D-2	87.54%
	D-3	86.94%
	D-4	86.82%
	Metrological reference	
	Max of T (D)	22.45%
	Avg I (D)	21.52%
	Max I (D)	21.03%
	Avg T (D-2)	21.17%
	Avg-Hi_opt	19.23%
	Avg-Hb(n)_m	19.36%
Month (M)	6.13%	
Input Model 03 (hourly)	D-1	88.83%
	D-2	87.54%
	Metrological reference	
	Max of T (D)	22.45%
	Avg I (D)	21.52%
	Max I (D)	21.03%
	Min T (D-1)	18.69%
	Avg T (D-2)	21.17%
	Avg-Hi_opt	19.23%
	Avg-Hb(n)_m	19.36%
	Month (M)	6.13%
Input Model 04 (hourly)	D-1	88.83%
	Metrological reference	
	Max of T (D)	22.45%
	Avg I (D)	21.52%
	Avg-Hb(n)_m	19.36%
Month (M)	6.13%	

In the second model of solar technology model, few of the variables were eliminated by retaining the high correlation value of each type variable. Here each type variable means, for example, MPH 01, D is historical value, Meteorological Value(T-Temperature M-month, I-Irradiation, H(h)_m: Irradiation on the horizontal plane, H(i_opt)_m: Irradiation on the optimally inclined plane, H(i)_m: Irradiation on a plane at an angle and Hb(n)_m: Monthly beam (direct) irradiation) and M-calender value month were all referred as type variable classify. From each classifies data with high correlation were retain the variable to the next models. data with the least correlation in the previous model will be eliminated. In this way, meticulous of production forecasting can be achieve with less error. In this second model, 2 historical variables and 3 meteorological variables were eliminated. variable from the different class was eliminated in this way the network stability and reduce the total error percentage. The variable data were taken into consideration is shortlisted in Table 24

In the third model 02 variable is eliminated from the previous model 02 and remaining data variable is considered as variable for model 03. Similarly in the fourth model05 variables were eliminated by which most high correlation of variable was only considered as input. In this way, the meticulous of production forecasting can be achieved. The variable which was considered is shortlisted in

Impending to the daily models it is pretty much similar to hourly models of Solar technology input variables entries considered most momentous in the enlargement of the historical energy production (D-1), (D-2) and (D-3) along with that Meteorological data variable such as Max of T (D), Avg I (D), Max I (D), Min T (D-1), Avg T(D-4), Avg-Hh, Avg-Hi_opt, Avg-Hi_m and Avg-Hb(n)_m and Month (M)with Pearson coefficient value of about 0.70, 0.66, 0.64 and Meteorological data Pearson coefficient 0.76, 0.73, 0.71, 0.63, 0.72, 0.62, 0.65, 0.64, 0.65 and 0.20 respectively.

Table 25 Solar technology Daily production forecasting models (MPD) entries variable with respective to correlation value in percentage

Types of Input Variable and respective Correlation Percentage			
Models (hourly)	Historical Production reference	Correlation Value in Percentage	
Input Model 01 (Daily)	D-1	70.84%	
	D-2	66.39%	
	D-3	64.47%	
	Metrological reference		
	Max of T (D)	76.34%	
	Avg I (D)	73.17%	
	Max I (D)	71.48%	
	Min T (D-1)	63.59%	
	Avg T (D-4)	72.01%	
	Avg-Hh	62.94%	
	Avg-Hi_opt	65.43%	
	Avg-Hi_m	64.86%	
	Avg-Hb(n)_m	65.90%	
	Month (M)	20.91%	
Input Model 02 (Daily)	D-1	70.84%	
	D-2	66.39%	
	Metrological reference		
	Max of T (D)	76.34%	
	Avg I (D)	73.17%	
	Max I (D)	71.48%	
	Avg T (D-4)	72.01%	
	Avg-Hh	62.94%	
	Avg-Hi_opt	65.43%	
	Avg-Hi_m	64.86%	
	Avg-Hb(n)_m	65.90%	
	Month (M)	20.91%	
	Input Model 03 (Daily)	D-1	70.84%
		Metrological reference	
Max of T (D)		76.34%	
Avg I (D)		73.17%	
Max I (D)		71.48%	
Avg T (D-4)		72.01%	
Avg-Hh		62.94%	
Avg-Hi_m		64.86%	
Avg-Hb(n)_m		65.90%	
Month (M)		20.91%	
Input Model 04 (Daily)	D-1	70.84%	
	Metrological reference		
	Max of T (D)	76.34%	
	Avg I (D)	73.17%	
	Avg-Hb(n)_m	65.90%	
Month (M)	20.91%		

Within the second model of the solar technology model, few of the variables were eliminated by holding the high correlation value of each sort of variable. From each classifies data with the high correlation were hold the variable to other models. Data with the slightest correlation within the previous model will be eliminated. In this way, fastidious production forecasting can be accomplished with fewer errors. In this second model01, historical variables and few meteorological variables were eliminated. variable from the different classes was eliminated in this way the network stability and diminish the whole error percentage. The variable data was taken into consideration is shortlisted in Table 25

Within the third model01 variable is eliminated from the previous model02 and the remaining data variable is considered as the variable for model03. Similarly within the fourth model03 variables were eliminated by which most high correlation of the variable was as it

was considered as input. In this way, meticulous production forecasting can be accomplished. The variable which was considered is shortlisted in Table 25.

8.2.2 Solar Technology Result and Discussion

After all the input variables were considered based on the above specification at that point the network is simulated particular output is generated. This output variable utilized as an input variable to calculate MAPE. The model that obtained a smaller MAPE in this Solar technology was MPH 02 with 38.00% within the hourly model and MPD 04 with 23.72% within the daily model which is shortlisted within the underneath table

Table 26 Solar Technology final model dropped values from

TECHNOLOGY	MODEL	MAPE
Solar(hourly)	Model 01	44.91%
	Model 02	38.00%
	Model 03	38.50%
	Model 04	49.97%
Solar(Daily)	Model 01	31.71%
	Model 02	25.24%
	Model 03	26.80%
	Model 04	23.72%

The mean value obtained for each input hour in its stimulation was utilized to test the model to verify its errors. The model which obtained lower error was considered as the best model for that specific technology. Each simulation was carried out considering the criteria of already decided with 5 output. Afterwards, the average output error obtained in the simulation resulted in MAPE. As expected the error value is higher than the previous model this justified since the load variable to forecast itself has related to the error. Be that as it may, the value remains underneath the MAPE of the best model making these the models selected for production forecast within the daily market.

Finally, was utilized as an explanation to proceed with the management of the error. Applying this process it was conceivable to obtain a reduction within the error of the ultimate model which is appeared in after all the input variables were considered based on the above specification at that point the network is simulated particular output is generated. This output variable utilized as an input variable to calculate MAPE. The model that obtained a smaller MAPE in this Solar technology was MPH 02 with 38.00% within the hourly model and MPD 04 with 23.72% within the daily model which is shortlisted within the underneath table

.Table 10 Summary of the method carried out to forecast the production of electricity within the development based on hourly and daily models

8.2.3 Solar Technology Result for Hourly Model

The correlation of the variable so obtained was not auspicious result as expected. This correlation of the variable was carried out by using the Pearson coefficient. Comparing with other models os hourly this obtained value of model no 02 is better than other models of hourly.

Then the obtained results of the forecast value were related to the actual production value along with test set variables. This information is graphically represented in Figure 34.

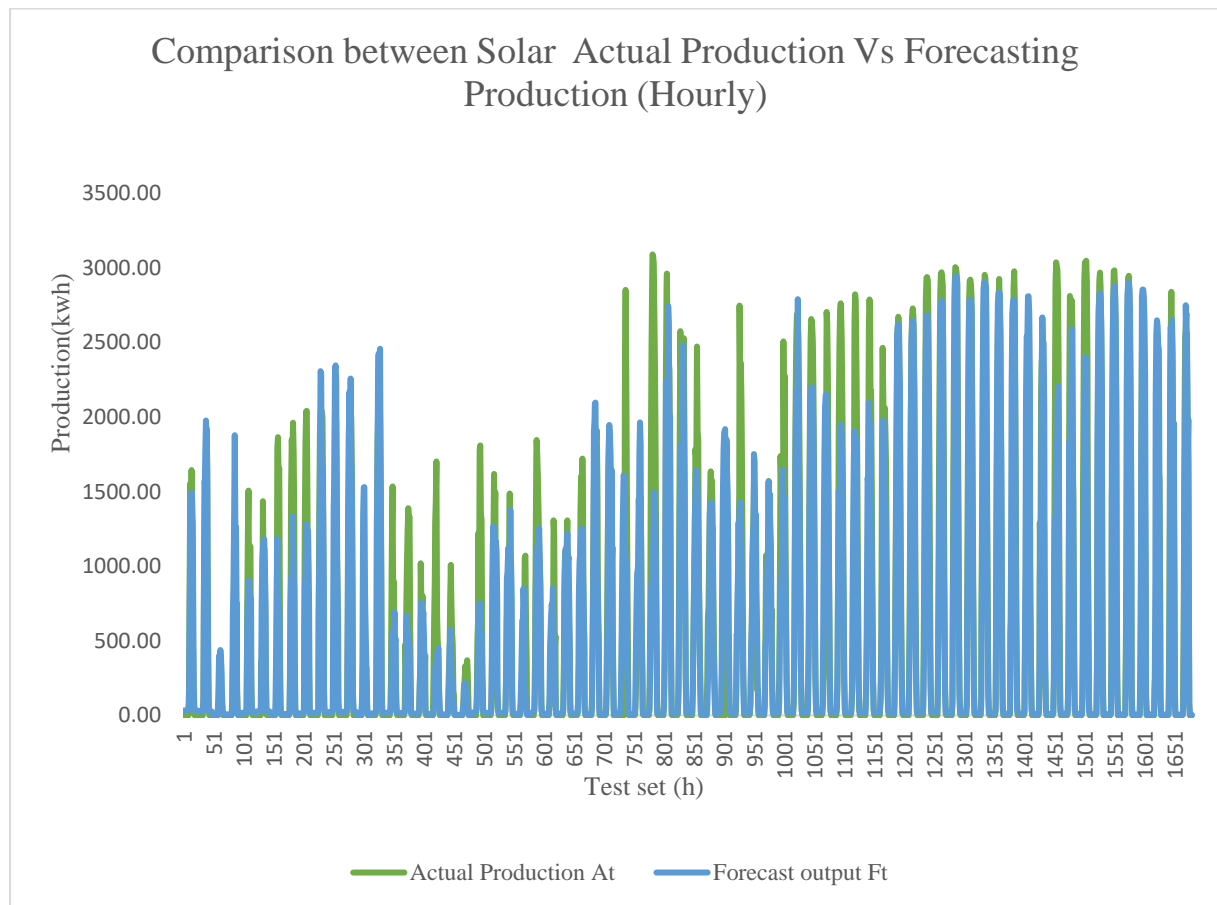


Figure 34 Comparison between the results obtained and actual production values (hourly model)

Figure 34 data about Solar technology production in Portugal additional time a period of a year (2019) summarises the data by selecting and reporting the most highlights to create comparisons that were relevant. Figure 26 for actual production, forecast production was nearly comparative in a certain time period and move distant absent in a certain time period at around 38.00 percent of error.

The actual production and forecast output conceivable outcomes frontier is developed by plotting all the conceivable combinations of the output represented in Figure 34. From the data represented within the graph, it is obvious that the test set point from 1035 to 1129 the two lines are a small a bit distant this has occurred when error distinguishes between actual and output production is higher. When it comes to testing set points 1 to 48 the line of the graph is near sufficient to one other this represents the forecasting variable at the point good with the smaller error were found.

The hourly model seems to have a greater correlation in following the Actual production when there are peaks once again the week of January 7 to 13 was isolated in 2018 with the results obtained within the simulation of this forecast model with a view to comparison with the values verified that week for electricity production Figure 35.

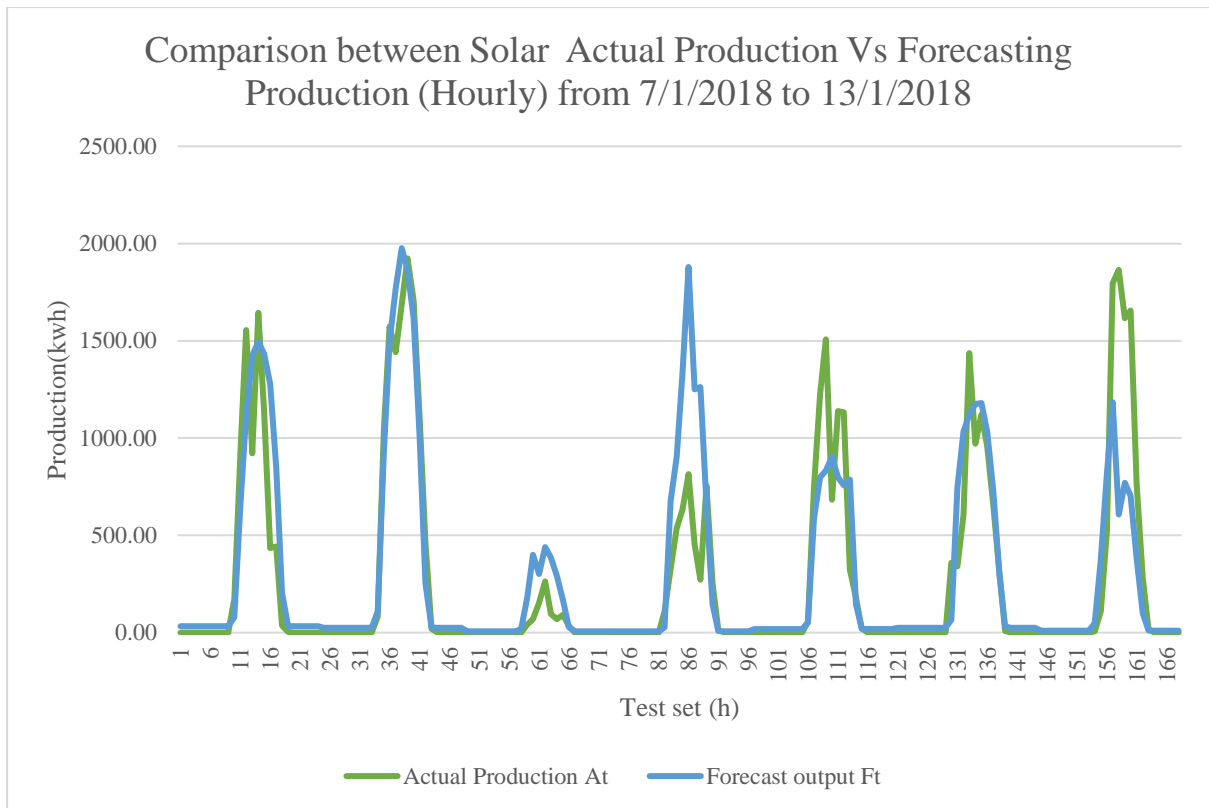


Figure 35 Week from January 7 to 13, 2018, with the comparison of the Actual values production and obtained outputs (Hourly model)

The Figure 35 information about Solar technology production in Portugal overtime a period of a year (2018) summarises the data by selecting and announcing the most highlights make comparisons where relevant. Here you can see the difference that exists between the outputs that are generated by the hourly model developed and the value analyzed alongside Actual energy production. As you can see the values produced by the model, usually slightly underneath the actual production values, with respect to the week that was chosen.

8.2.4 Solar Technology Obtained Error for Hourly Model

The histogram chart represents the error difference between the actual production and simulated 5 output avg forecasting production of the hourly model. This graph gives data about the colossal of average error confirming with absolute production value. when the deviation of value between the average forecast value and actual production is high this leads to an error event this analysis of the data has appeared within the Figure 36, Figure 37. Represented the histogram of the relatively obtained output production error can be established

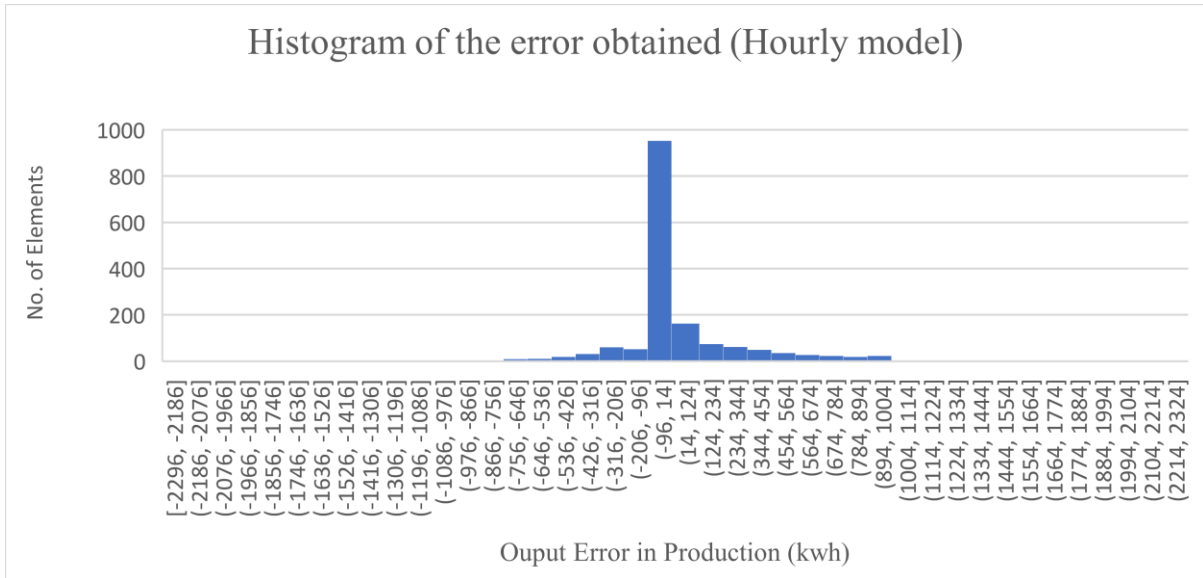


Figure 36 Example Histogram of the Production error obtained (Hourly model)

The graph shows information about production errors obtained will creating and executing the model centred on the interval containing zero and are relatively distributed around these values. It represents relative error over the sample of the test set used in Figure 37

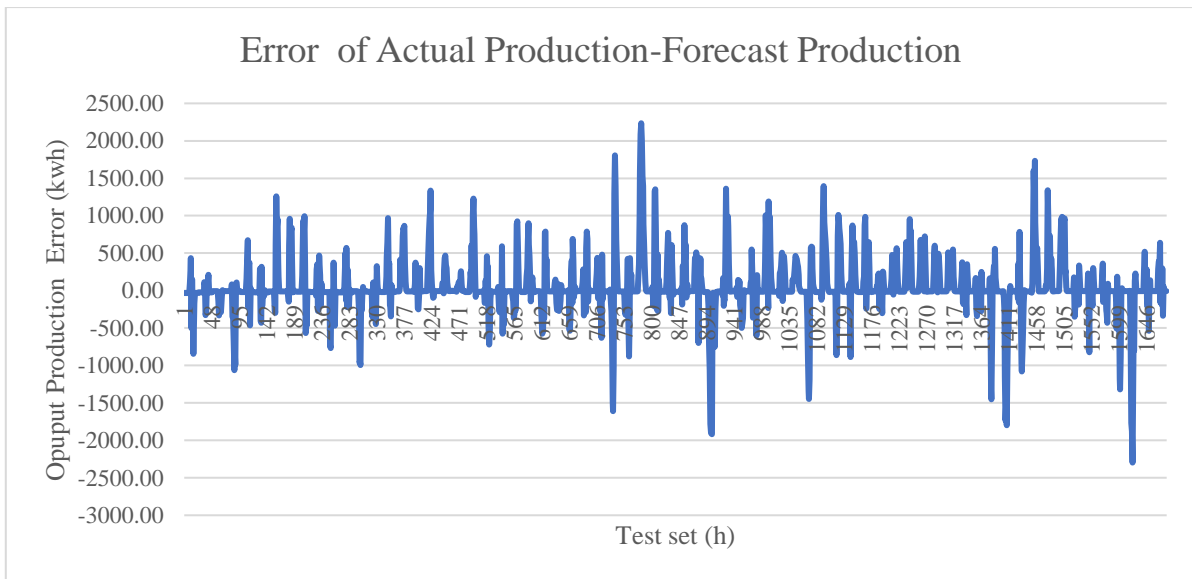


Figure 37 Example Obtained output production relative error and standard deviation (Hourly model)

The Figure 37 graphical representation gives information about the highest and lowest error obtained from output production. These variables are compared with the hourly test set variable. the error of actual – output production is the peaks that occur are represented as peaks in the graph. The colossal error found was around -2296.14 kWh and may be due to the great flaw pointed to the method used for the performance calculation of the MAPE. This collapse is due to obtaining Actual- forecast output production values of zero or close to zero which make the error found to be very high regardless of the error absolute value. The results obtained and calculations performed in this model can be seen in Annex2.

8.2.5 Solar Technology Result for Daily Model

The correlation of the variable so obtained was not auspicious results as expected. This correlation of the variable was carried out by using the Pearson coefficient. Comparing with other models os hourly this obtained value of model no 04 is better than other models of hourly. Then the obtained results of the forecast value were related to the actual production value along with test set variables. This information is graphically represented in Figure 38. However, the values are able to follow their evolution despite its volatility. This can be seen in the daily model the difference that exists between the outputs that are generated in the hourly model developed is comparable small then Daily model. Daily model representation graph shown beneath with the average MAPE verified for this model in the test set

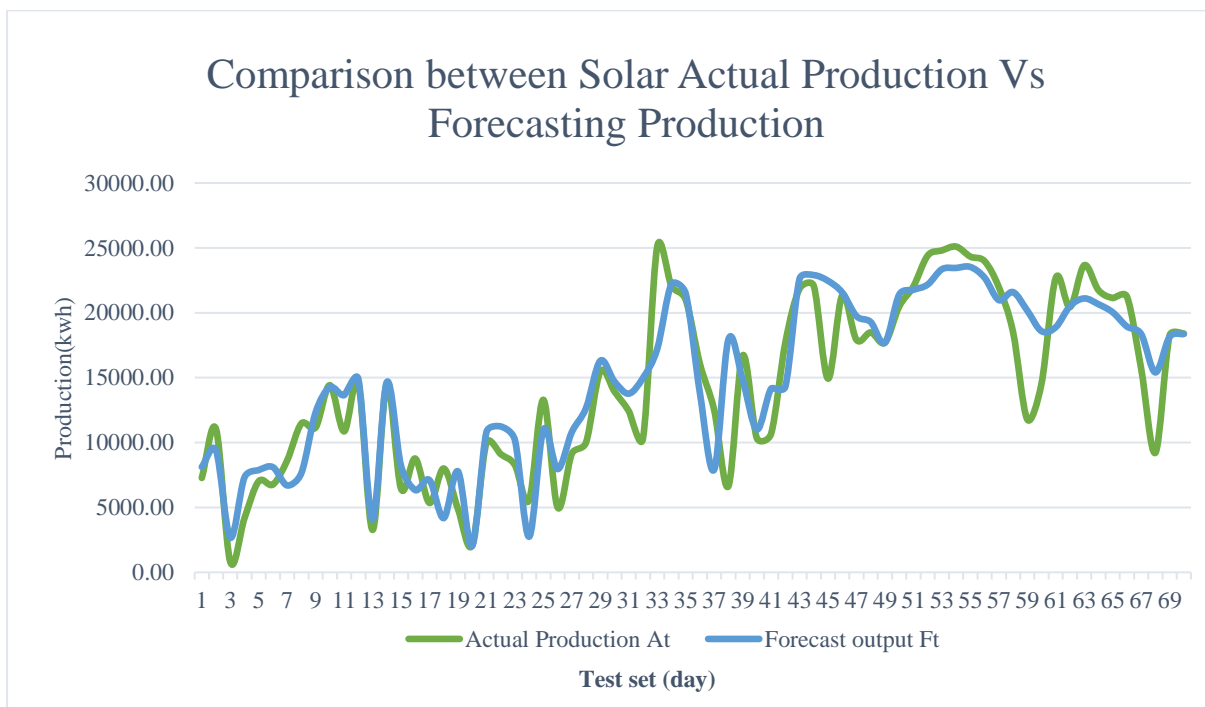


Figure 38 Comparison between the results obtained and actual production values (Daily model)

Figure 38 symbolize information about Solar technology production in Portugal overtime period a year (2019) summarise the information by selecting and reporting the main features to make comparisons where relevant. Figure 38. for actual production, forecast production was almost similar in the certain time period and move far away in a certain time period at around 23.72 percent of error.

The actual production and forecast output possibilities frontier is constructed by plotting all the possible combination of the output represented in Figure 36. From the information represented in the graph it is clear that the test set point at 31 to 35 the two lines is a little bit distant this has occurred when error distinguishes between actual and output production is higher. When it comes to testing set point 13 to 15 the line of the graph are close enough to one other this represents the forecasting variable at the point good with the smaller error were found.

The Daily model it seems to have a greater correlation in following the Actual production when there are peaks once again the week of January 6 to 12 was isolated 2019 with

the results obtained in the simulation of this forecast model with a view to comparison with the values verified that week for electricity production Figure 39.

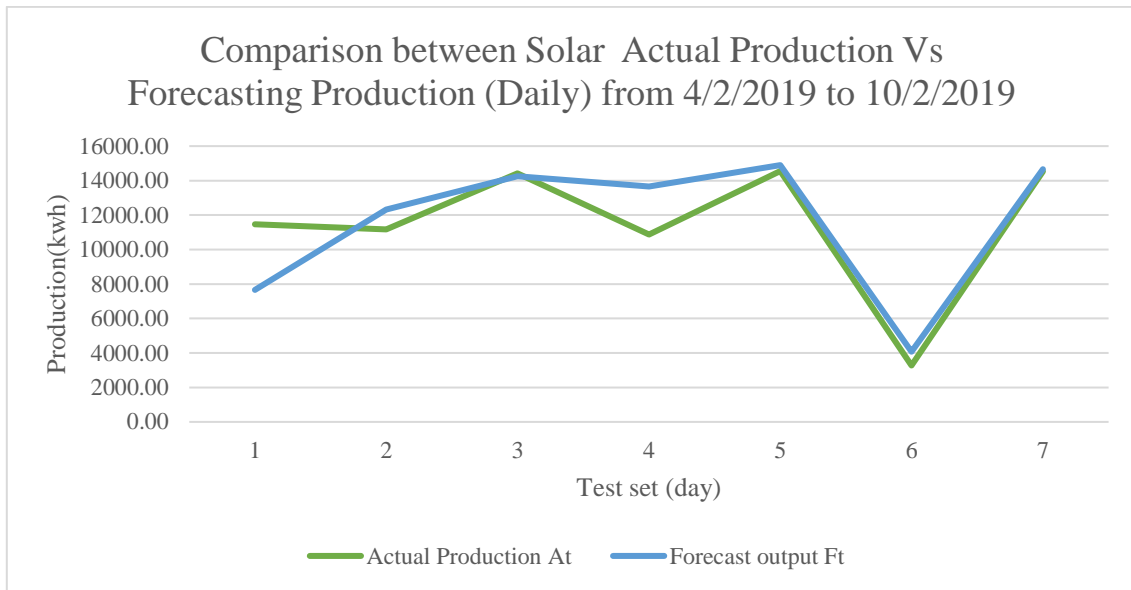


Figure 39 Week from February 4 to 10 (2018), with the comparison of the Actual values of production and obtained outputs (Daily model)

The Figure 39 information about Solar technology production in Portugal overtime period a year (2018) summarise the information by selecting and reporting the main features to make comparisons where relevant.

Here you can see the difference that exists between the outputs that are generated by the Daily model developed and the value analyzed along with Actual energy production. As you can see the values generated by the model, usually slightly below the actual production values, with respect to the week that was chosen.

8.2.6 Solar Technology Obtained Error for Daily Model

The histogram graph represents the error difference between the actual production and simulated 5 output avg forecasting production of the hourly model. This graph gives information about colossal of average error confirming with absolute production value. when the deviation of value between the average forecast value and actual production is high this leads to error occurrence this analysis of the data is shown in the below Figure 40,Figure 41. Represented the histogram of the relative obtained output production error can be established.

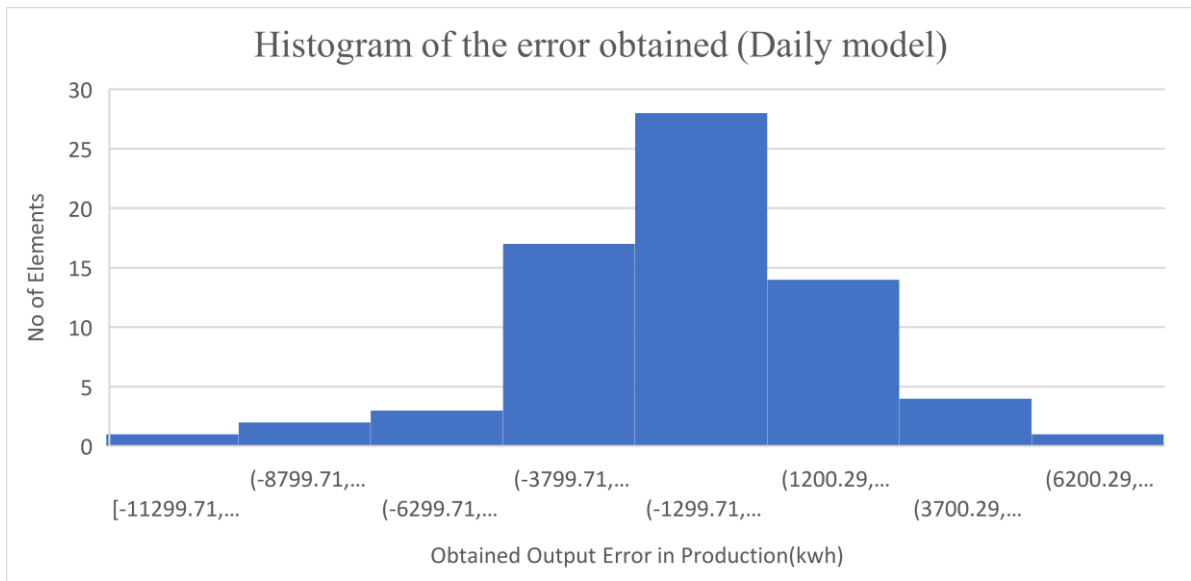


Figure 40 Example Histogram of the Production error obtained (Daily model)

The graph shows information about production errors obtained will creating and executing the model centred on the interval containing zero and are relatively distributed around these values. It represents relative error over the sample of the test set used in Figure 41.

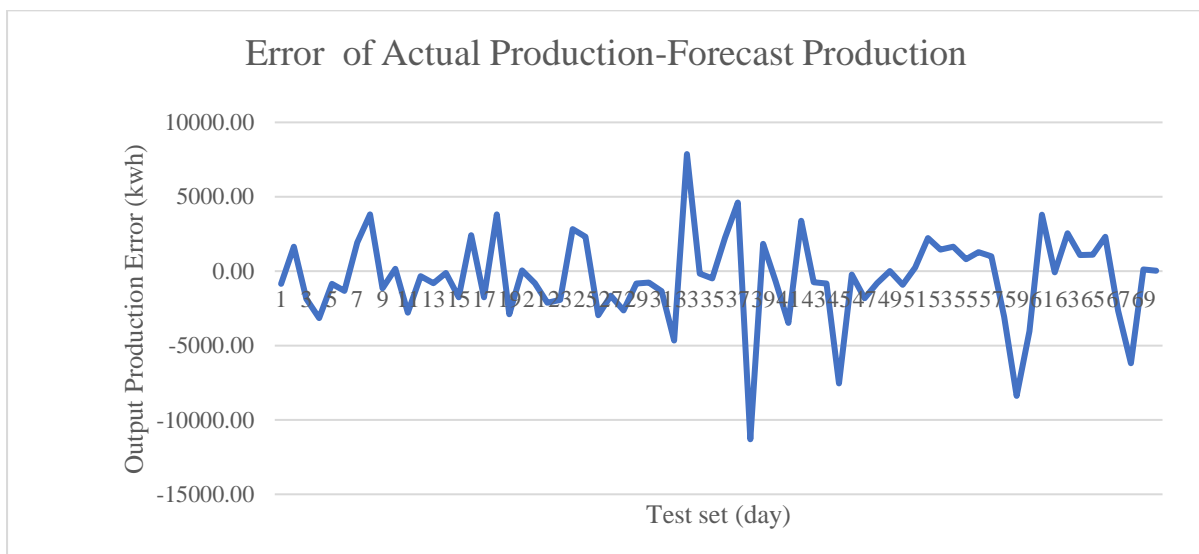


Figure 41 Example Obtained Output Production relative error and standard deviation (Daily model)

The Figure 41 graphical representation gives information about the highest and lowest error obtained from output production. These variables are compared with the hourly test set variable. the error of actual – output production is the peaks that occur are represented peaks in the graph. The colossal error found was around -11299.70 kWh and may be due to the great flaw pointed to the method used for the performance calculation of the MAPE. This collapse is due to obtaining Actual- forecast output production values of zero or close to zero which make the error found to be very high regardless of the error absolute value. The results obtained and calculations performed in this model can be seen in the Annex2.

8.3 Wind Technology

In the case of wind technology, the variable such as Historical data and Meteorological data were gathered from the different platform from which variables with high correlation is considered as input. Later tested multiple times to reduces the false-positivity of data set. In this way, data sets were well organized. This correlation of variables was calculated by person coefficient. Input variables identified for each model differ from one other. This variable was carried out to test multiple times which lead forward to attain least error forecasting models.

8.3.1 Wind Technology Input Variable

For the first model of wind technology hourly the variables identified and considered most momentous in the enlargement of the historical energy production (D-1), along with that Meteorological data variable such as Avg WS(D) and MinWS(D) with a Pearson coefficient value of about 0.35 Meteorological data Pearson coefficient 0.60 and 0.58 respectively. This all variable was divided 80:20 ratio as it briefly explained in the above section in this 80% is input and 20% is the samples to be tested. This classification is respected for each and every model along with there respective technology.

One of the most significant things is that while considering the variables of previous historical and pervious Meteorological data most behind variable with high correlation data input is considered from that and beneath data would be taken as consideration above value would be eliminated. For example, MPD 01 consist of the different high correlation of variable were considered as input in that least historical or most previous Meteorological data, previous day-6 value are higher still the value is selected from 7 days and beneath in this way the data variable was selected

Table 27 Wind technology Hourly production forecasting models (MPH) entries variable with respective to correlation value in percentage

Types of Input Variable and respective Correlation Percentage		
Models (hourly)	Historical Production reference	Correlation Value in Percentage
Input Model 01 (hourly)	D-1	35.82%
	Metrological reference	
	Avg WS (D)	60.11%
	Min WS (D)	58.42%
Input Model 02 (hourly)	D-1	35.82%
	Metrological reference	
	Avg WS (D)	60.11%
	Month (M)	
Input Model 03 (hourly)	D	100.00%
	Metrological reference	
	Avg WS (D)	60.11%

In the second model of the Wind Technology model, few of the variables were eliminated or eliminated by retaining the high correlation value of each type variable. Here each type variable means, for example, MPH 01, D is historical value, Meteorological Value(Average Wind speed (Avg WS) Minium Wind speed (Min WS) were all referred as type variable classify. From each classifies data with high correlation were retain the variable to the next models. Data with the least correlation in the previous model will be eliminated. In this way, meticulous of production forecasting can be achieved with less error. In this second

model01 Meteorological variable was eliminated. variable from the different class was eliminated in this way the network stability and reduce the total error percentage. The variable data were taken into consideration is shortlisted in Table 27

In the third model01 variable is eliminated in this case historical data and 1 variable is added which is the actual production value of D from the previous model 02 and remaining data variable is considered as a variable for model 03. In this way, the meticulous of production forecasting can be achieved. The variable which was considered is shortlisted in Table 27

Approaching to the daily models it is pretty much similar to hourly models of Wind technology input variables entries considered most momentous in the enlargement of the historical energy production (D-1),(D-2) and (D-3) along with that Meteorological data variable such as Max of T (D), Avg I (D), Max I (D), Min T (D-1), Avg T (D-4), Avg-Hh, Avg-Hi_opt, Avg-Hi_m and Avg-Hb(n)_m and Month (M)with a Pearson coefficient value of about 0.70,0.66,0.64 and Meteorological data Pearson coefficient 0.76, 0.73, 0.71, 0.63, 0.72, 0.62, 0.65, 0.64, 0.65 and 0.20 respectively.

Table 28 Wind technology Daily production forecasting models (MPD) entries variable with respective to correlation value in percentage

Types of Input Variable and respective Correlation Percentage		
Models (hourly)	Production reference	Correlation Value in Percentage
Input Model 01 (Daily)	D-1	47.85%
	Metrological reference	
	Avg WS (D)	77.12%
	Min WS (D)	74.96%
	Max of WS (D)	63.69%
	Avg WS (D-1)	42.98%
	Min WS (D-1)	44.90%
Input Model 02 (Daily)	D-1	47.85%
	Metrological reference	
	Avg WS (D)	77.12%
	Min WS (D)	74.96%
	Max of WS (D)	63.69%
Input Model 02 (Daily)	D-1	47.85%
	Metrological reference	
	Avg WS (D)	77.12%
	Min WS (D)	74.96%
	Max of WS (D)	63.69%
Input Model 03 (Daily)	D-1	47.85%
	Metrological reference	
	Avg WS (D)	77.12%
Input Model 04 (Daily)	D-1	47.85%
	Metrological reference	
	Avg WS (D)	77.12%
Input Model 05 (Daily)	D	100.00%
	Metrological reference	
	Avg WS (D)	77.12%

In the second model of the Wind Technology model, few of the variables were eliminated by retaining the high correlation value of each type variable. From each classifies data with high correlation were retain the variable to the next models. Data with the least correlation in the previous model will be eliminated. In this way, meticulous of production forecasting can be achieved with less error. In this second model01, historical variables and 2

Meteorological variables were eliminated. variable from the different class was eliminated in this way the network stability and reduce the total error percentage. The variable data were taken into consideration is shortlisted in the Table 28

In the third model01 variable is eliminated from the previous model 02 and remaining data variable is considered as variable for model 03. Similarly in the fourth model03 variables were eliminated by which most high correlation of variable was only considered as input. In this way, the meticulous of production forecasting can be achieved. The variable which was considered is shortlisted in Table 28.

8.3.2 Wind Technology Result and Discussion

After all the input variable were consider based on the above specification then the network is simulated respective output is generated. This output variable used as the input variable to calculate MAPE. The model that obtained a smaller MAPE in this Wind technology was MPH 01 with 480.29% in the hourly model and in MPD 03 with 55.31% in the daily model which is shortlisted in the below table

Table 29Wind Technology final model dropped values from

TECHNOLOGY	MODEL	MAPE
Wind(hourly)	Model 01	480.29%
	Model 02	984.40%
	Model 03	1161.05%
Wind(Daily)	Model 01	63.63%
	Model 02	65.74%
	Model 03	55.31%
	Model 04	58.37%
	Model 05	75.18%

The mean value obtained for each input hour in its stimulation were used to test the model to verify its errors. the model which obtained lower error was considered as the best model for that particular technology. Each simulation was carried out considering the criteria of previously determined with 5 output. Later average output error obtained is simulation resulted from MAPE. As expected the error value is higher than the previous model this justified since the load variable to forecast itself has associated with the error. However, the value remains below the MAPE of the best model making these the models selected for production forecast in the daily market. Finally, it was used explanation to proceed with management of the error.

Applying this process it was possible to obtain a reduction in the error of the final model which is shown in after all the input variable were consider based on the above specification then the network is simulated respective output is generated. This output variable used as the input variable to calculate MAPE. The model that obtained a smaller MAPE in this Wind technology was MPH 01 with 480.29% in the hourly model and in MPD 03 with 55.31% in the daily model which is shortlisted in the below table

Table 10 Summary of the method carried out to forecast the production of electricity in the development based on hourly and daily models.

8.3.3 Wind Technology Result for Hourly Model

The correlation of the variable so obtained was not auspicious result as expected. This correlation of the variable was carried out by using the Pearson coefficient. Comparing with other models os hourly this obtained value of model no 01 is better than other models of hourly. Then the obtained results of the forecast value were related to the actual production value along with test set variables. This information is graphically represented in the Figure 42.

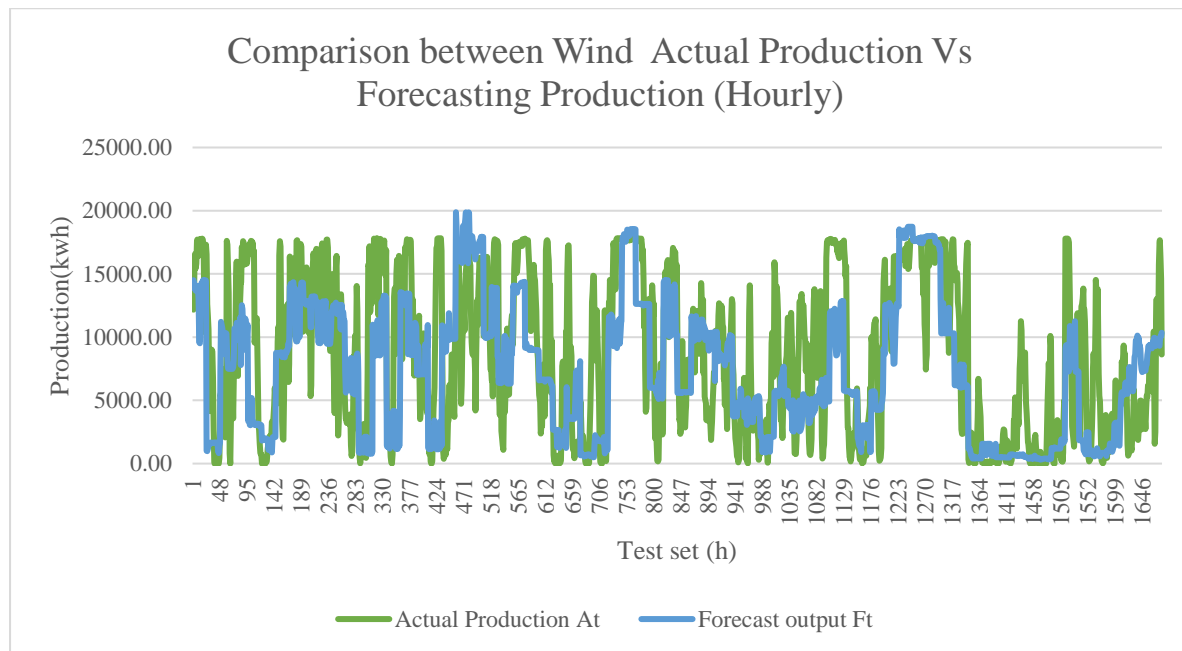


Figure 42 Comparison between the results obtained and actual production values (hourly model)

The Figure 42 information about Wind technology production in Portugal overtime period a year (2018) summarise the information by selecting and reporting the main features make comparisons where relevant. Figure 26 for actual production, forecast production was almost similar in the certain time period and move far away in a certain time period at around 480.29 percent of error.

The actual production and forecast output possibilities frontier is constructed by plotting all the possible combination of the output represented in Figure 42. From the information represented in the graph the test set point at 193 to 389 the two lines is a little bit distant this has occurred when error distinguishes between actual and output production is higher. When it comes to testing set point 1 to 49 the line of the graph is close enough to one other this represents the forecasting variable at the point good with the smaller error were found.

The hourly model it seems to have a greater correlation in following the Actual production when there are peaks once again the week of January 7 to 13 was isolated 2018 with the results obtained in the simulation of this forecast model with a view to comparison with the values verified that week for electricity production Figure 43

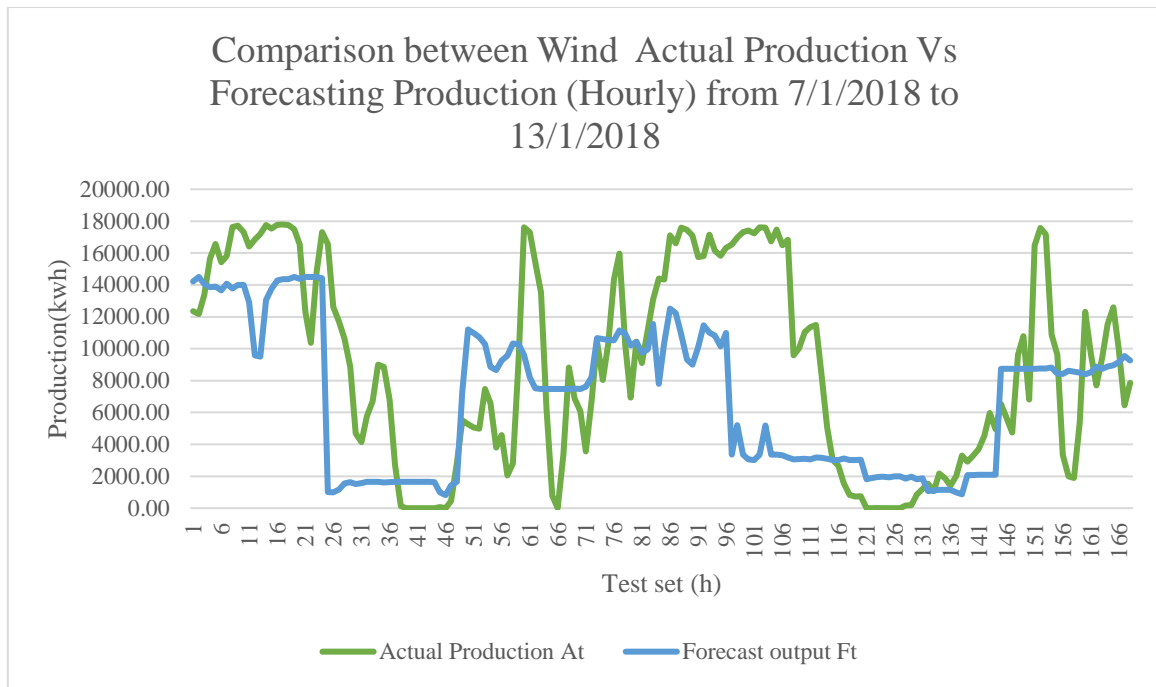


Figure 43 Week from January 7 to 13, 2018, with the comparison of the Actual values of production and obtained outputs (Hourly model)

The Figure 43 information about Wind technology production in Portugal overtime period a year (2018) summarise the information by selecting and reporting the main features make comparisons where relevant.

Here you can see the difference that exists between the outputs that are generated by the hourly model developed and the value analyzed along with Actual energy production. As you can see the values generated by the model, usually slightly below the actual production values, with respect to the week that was chosen.

8.3.4 Wind Technology Obtained Error for Hourly Model

The histogram graph represents the error difference between the actual production and simulated 5 output avg forecasting production of the hourly model. This graph gives information about colossal of average error confirming with absolute production value. when the deviation of value between the average forecast value and actual production is high this leads to error occurrence this analysis of the data is shown in the below Figure 44 Figure 45. Represented the histogram of the relative obtained output production errors can be established.

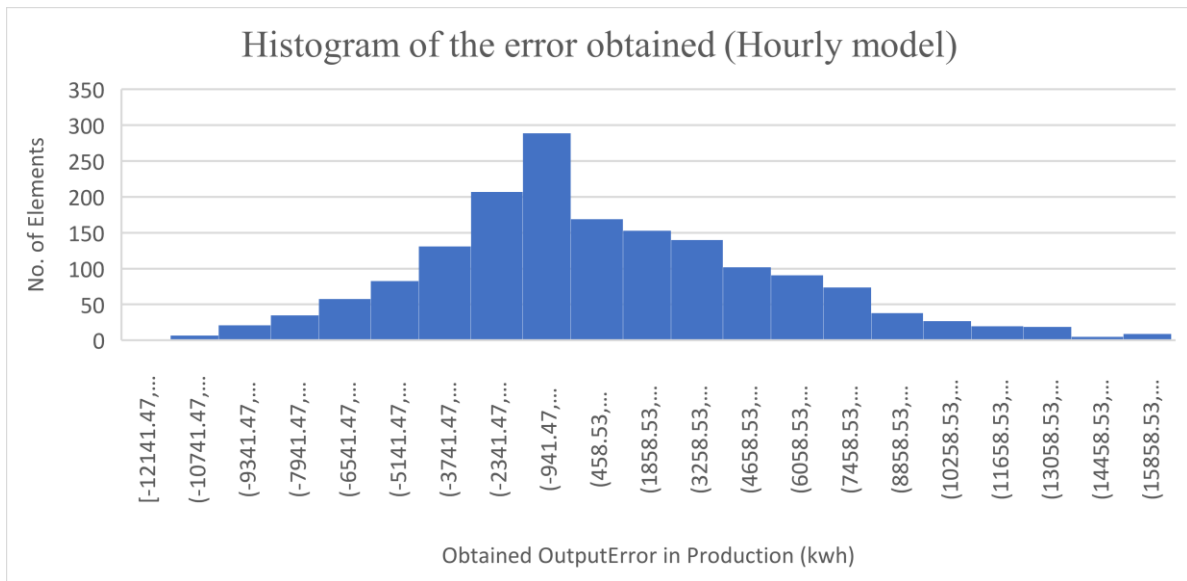


Figure 44 Example Histogram of the Production error obtained (Hourly model)

The graph shows information about production errors obtained will creating and executing the model centred on the interval containing zero and are relatively distributed around these values. It represents relative error over the sample of the test set used in Figure 45.

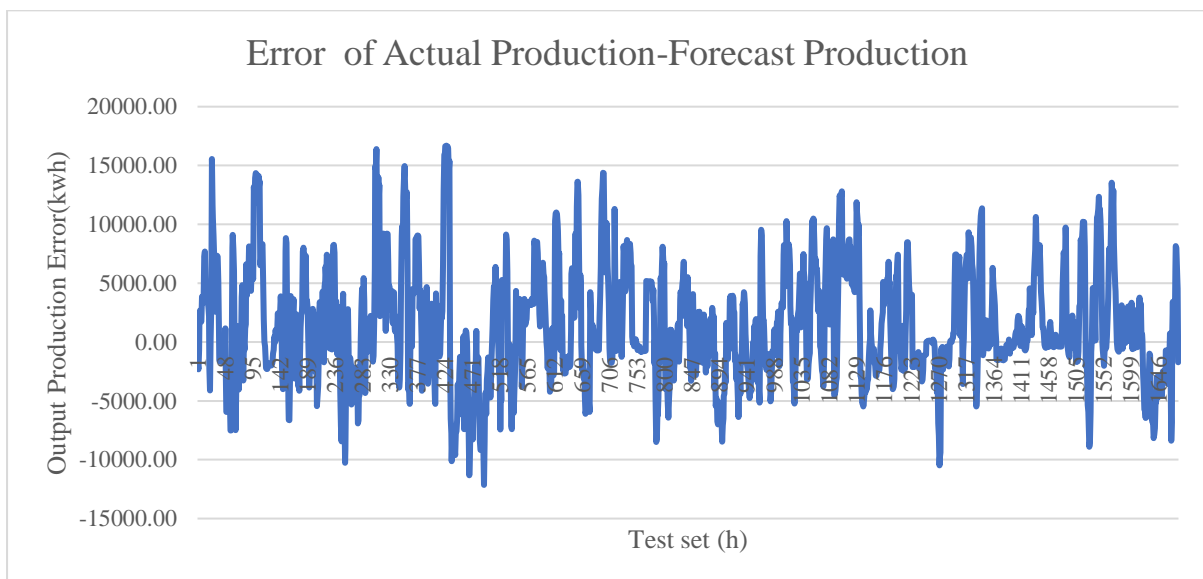


Figure 45 Example Obtained output production relative error and standard deviation (Hourly model)

Figure 45 graphical representation gives information about the highest and lowest error obtained from output production. These variables are compared with the hourly test set variable. the error of actual – output production is the peaks that occur are represented peaks in the graph. The colossal error found was around -12141.46 kWh and may be due to the great flaw pointed to the method used for the performance calculation of the MAPE. This collapse is due to obtaining Actual- forecast output production values of zero or close to zero which make the error found to be very high regardless of the error absolute value. The results obtained and calculations performed in this model can be seen in the Annex3

8.3.5 Wind Technology Result for Daily Model

The correlation of the variable so obtained was not auspicious result as expected. This correlation of the variable was carried out by using the Pearson coefficient. Comparing with other models os hourly this obtained value of model no 03 is better than other models of hourly. Then the obtained results of the forecast value were related to the actual production value along with test set variables. This information is graphically represented in the Figure 46. However, the values are able to follow their evolution despite its volatility. This can be seen in the daily model the difference that exists between the outputs that are generated in the hourly model developed is comparable small then Daily model. Daily model representation graph shown beneath with the average MAPE verified for this model in the test set.

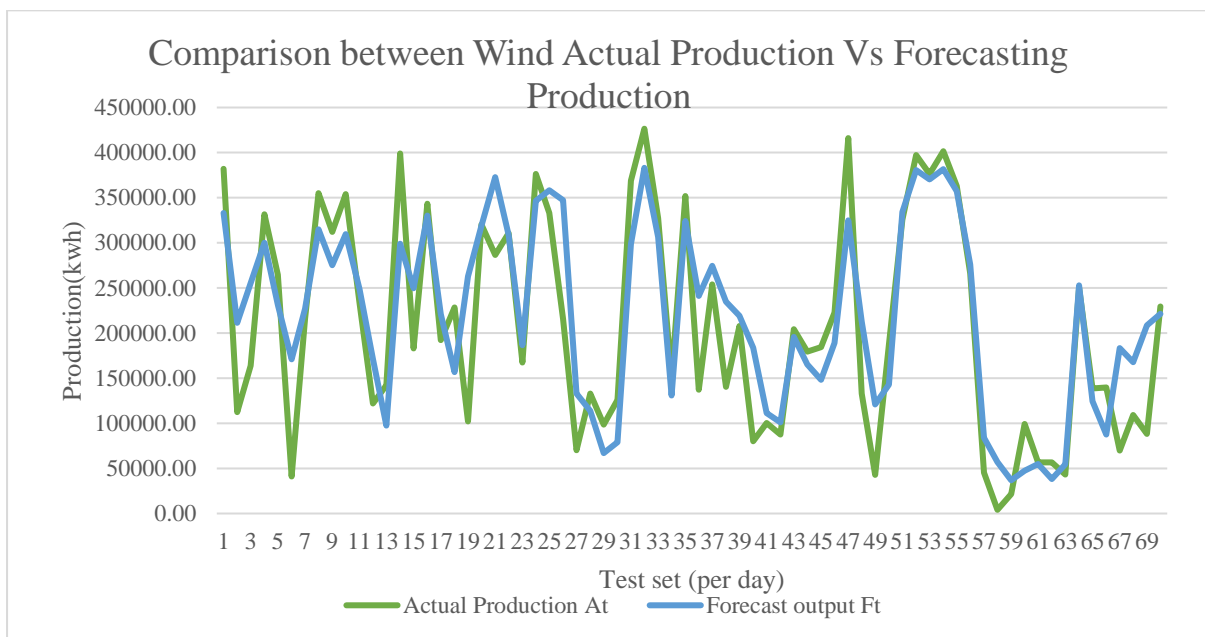


Figure 46 Comparison between the results obtained and actual production values (Daily model)

The Figure 46 symbolize information about Wind technology production in Portugal overtime period a year (2019) summarise the information by selecting and reporting the main features to make comparisons where relevant. figure 44. for actual production, forecast production was almost similar in the certain time period and move far away in a certain time period at around 23.72 percent of error.

The actual production and forecast output possibilities frontier is constructed by plotting all the possible combination of the output represented in Figure 46. From the information represented in the graph it is clear that the test set point at 23 to 25 the two lines is a little bit distant this has occurred when error distinguishes between actual and output production is higher. When it comes to testing set point 1 to 7 the line of the graph are close enough to one other this represents the forecasting variable at the point good with the smaller error were found.

The Daily model it seems to have a greater correlation in following the Actual production when there are peaks once again the week of January 6 to 12 was isolated 2019 with the results obtained in the simulation of this forecast model with a view to comparison with the values verified that week for electricity production Figure 47.

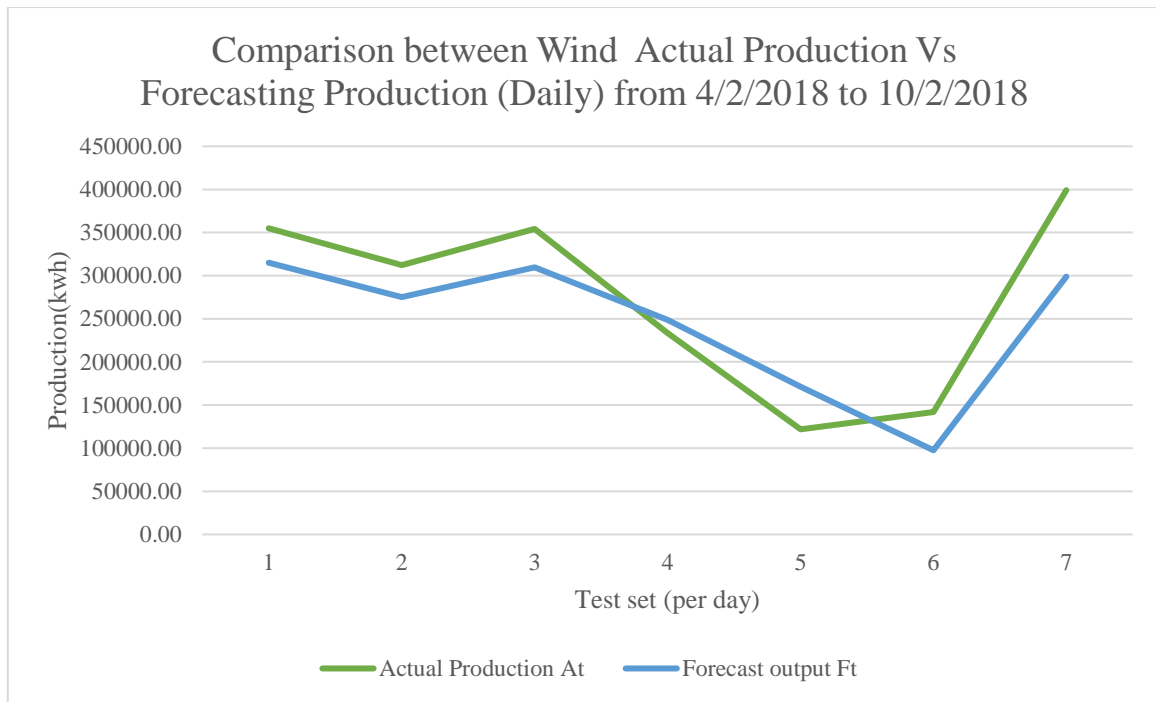


Figure 47 Week from February 4 to 10 (2018), with the comparison of the Actual values of production and obtained outputs (Daily model)

The Figure 47 information about Wind technology production in Portugal overtime period a year (2018) summarise the information by selecting and reporting the main features make comparisons where relevant.

Here you can see the difference that exists between the outputs that are generated by the Daily model developed and the value analyzed along with Actual energy production. As you can see the values generated by the model, usually slightly below the actual production values, with respect to the week that was chosen.

8.3.6 Wind Technology Obtained Error for Daily Model

The histogram graph represents the error difference between the actual production and simulated 5 output avg forecasting production of the hourly model. This graph gives information about colossal of average error confirming with absolute production value. when the deviation of value between the average forecast value and actual production is high this leads to error occurrence this analysis of the data is shown in the below Figure 48, Figure 49. Represented the histogram of the relative obtained output production errors can be established.

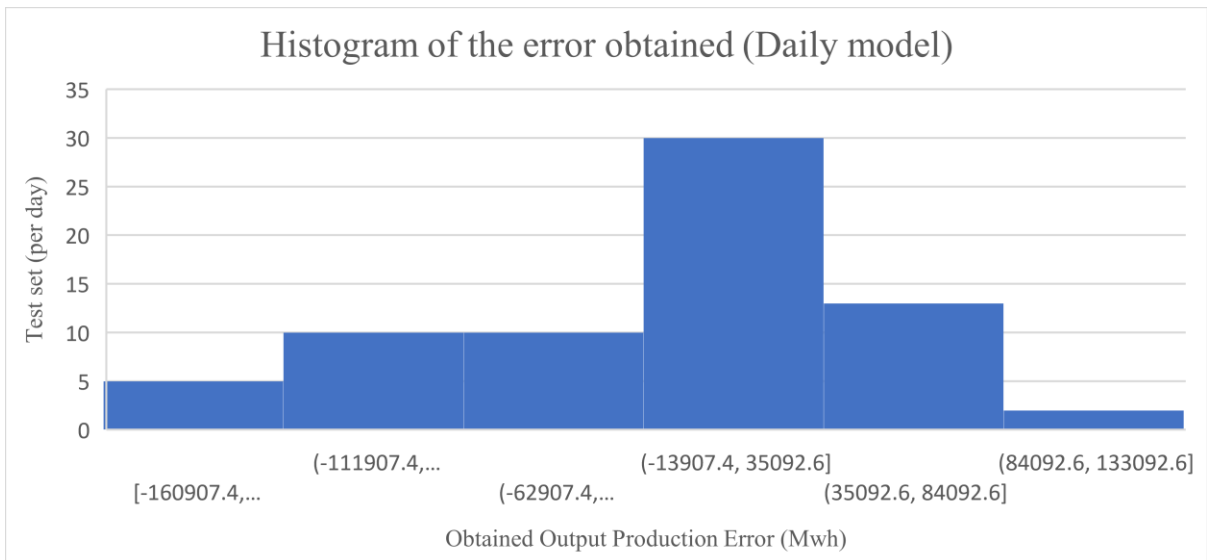


Figure 48 Example Histogram of the Production error obtained (Daily model)

The graph shows information about production errors obtained will creating and executing the model centred on the interval containing zero and are relatively distributed around these values. It represents relative error over the sample of the test set used in Figure 49.

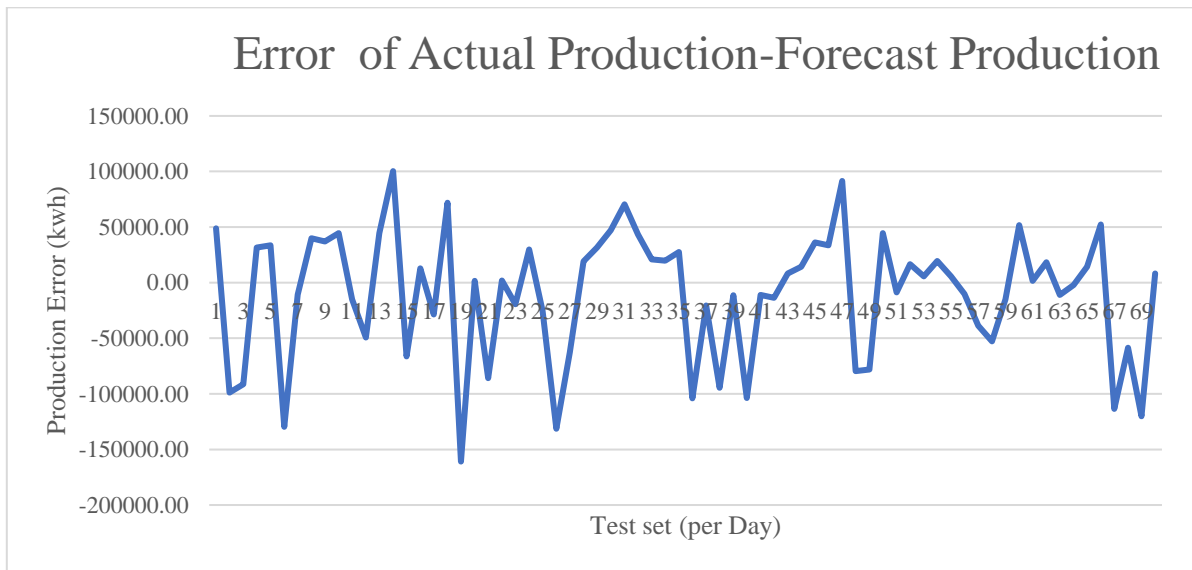


Figure 49 Example Obtained output production relative error and standard deviation (Daily model)

The Figure 49 graphical representation gives information about the highest and lowest error obtained from output production. These variables are compared with the hourly test set variable. the error of actual – output production is the peaks that occur are represented peaks in the graph. The colossal error found was around -160907.38 kWh and may be due to the great flaw pointed to the method used for the performance calculation of the MAPE. This collapse is due to obtaining Actual- forecast output production values of zero or close to zero which make the error found to be very high regardless of the error absolute value. The results obtained and calculations performed in this model can be seen in the Annex3.

8.4 Risk Magement control for Hydro Technology

The colossal error found was around -4811.6 kWh (hourly model)-76488.08 kWh (daily model) for hydro technology, colossal error found was around -2296.14 kWh (hourly model) -11299.70 kWh (daily model) for solar technology and colossal error found was around -12141.46 kWh (hourly model) -160907.38 kWh (daily model) for wind technology in arranging to overcome this demand in a balance circumstance ideal the wind turbines can be utilized to boost energy production and can compensate for the demand. According to a few research, organizations seem to have reinforcement ideal technology to bolster as a reinforcement. Another way it can be boosted in other technologies based on meteorological factors in any case high of irradiation on that specific day at that point few solar panels can be joined to compensate demand. Possibly pluviosity is higher due to bounty of precipitation and the reservoir is adequate for the running of the hydro turbine may offer assistance production this may only be done to adjust demand and supply. By following this way the organization can get a small benefit but this can be profoundly dependent on the spot market price. In the event that any case demand is over-burdening, it may affect the spot market price simultaneously it will influence the production it is the accentuation to preserve the demand and supply. This dissertation makes a difference in the organization to choose when to charge energy production and when to back by keeping up demand and supply.

8.5 Section Summary.

This outstanding way of studying the method for forecasting production variables was tested with the assorted variable. This variable also tested with the correlation of assorted variables. Data variable with highest is considered as input and sample variable.

These variables were used to allows the construction of nntool is part of the set of tools provided by the Matlab ® software. By using this nntool were selected and then tested were carried out for each time period with an assorted combination of variables and number of training sessions with the objective of improving the results.

Based on the above results on technology-wise obtained the variables that should be included in the models to be created for each time period in order to guarantee the best possible result, several models have been developed for each scenario, using variable have a better correlation. In the hourly and daily models, this process was tested has been shown to achieve better results compared to previous models.

Then, the results that were obtained in the models selected and were subject to an error management process with the aspiration of accomplishing the minium error this process was based on the calculation of the standard deviation error of the final outputs of models selected in each case, the output of each stimulation corresponding to each time period hourly and daily models this resulting value would then be the one that would be considered. In the case of the daily model, the value of APE accomplish despite being close to what had been obtained prior to error management, is slightly higher. In few cases likewise, MPH 01, MPH02 and MPH03 of wind technology in all this the values of error got lousy.

9 CONCLUSIONS

9.1 Final Conclusions And Main Contributions

This dissertation primary focus to develop models forecast of electricity production, suitable for competitive markets, utilizing the tools necessary. The tool is chosen after a detailed study was nntool, beneath the Artificial Neural Networks (ANN) method in Matlab ® software. In the beginning, it is fundamental to gather a large number of data that was considered to have an influence on the evolution of production values its imperative to gather a number of doable data, paying attention to its quality of data variable. This data was afterwards tested and considered in order to scrutinize be the foremost crucial step for the training of the ANNs is developed. This may be utilized as inputs to adjust the information with the time period for each model developed.

Regardless of the time period, the method by which it was chosen to carry out the proposed objective, was through the development of the hourly and daily models with the distinctive variables accessible attempting to discover the most excellent combination by which the error obtained surveyed through MAPE was minimum with this the values of the parameters were kept constant throughout the work of developing the network. After making all the network models along with the method of testing by training numerous times before simulate. Once simulate is done the output is utilized within the error management system to calculate MAPE.

After checking the MAPE of the forecasting model the predicted value after treatment was utilized as input to the fruitful model of the production forecast. When the fresh forecast is made, initially contains a more prominent error than already generated. Be that as it may, in order to achieve minimum error best model of each technology would be chosen as the final models. this occurred for the hourly and daily time period being that as it were in a model in each technology with input that got error smaller.

Then by calculating the standard deviation of the outputs along with time period and MAPE were calculated. Finally, the results obtained were analyzed for each of the developed models of production forecast. In order to fathom characteristics of the error obtained in each case. This was done using the study of the Pearson Coefficient output compared to actual production values. Using the graphical demonstration of the relationship of the two variables and later represented in graphically, some other represented were relatively obtained output production errors obtained and a histogram showing where the midpoint of the error predictions in each model.

This all conclude the obtained results of one model hourly and one model were decent and good to some extent indicative of the production rest of all models in each technology.

The construction of the models that obtained the best results were shortlisted beneath.

Technology	Best Model Results	Avg.MAPE	Network Training	MAPE
Hydro(hourly)	Model 02	38.84%	Network 1	47.22%
			Network 2	46.61%
			Network 3	46.70%
			Network 4	39.70%
			Network 5	38.61%
Hydro(Daily)	Model 03	21.44%	Network 1	25.68%
			Network 2	21.97%
			Network 3	26.79%
			Network 4	24.79%
			Network 5	26.33%
Solar(hourly)	Model 02	38.00%	Network 1	49.64%
			Network 2	47.12%
			Network 3	46.30%
			Network 4	46.82%
			Network 5	50.44%
Solar(Daily)	Model 04	23.72%	Network 1	26.54%
			Network 2	24.68%
			Network 3	24.37%
			Network 4	26.18%
			Network 5	25.95%
Wind(hourly)	Model 01	480.29%	Network 1	363.87%
			Network 2	263.92%
			Network 3	300.07%
			Network 4	519.32%
			Network 5	1083.38%
Wind(Daily)	Model 03	55.31%	Network 1	56.67%
			Network 2	62.52%
			Network 3	51.58%
			Network 4	56.76%
			Network 5	55.36%

9.2 Future Work

During this dissertation study and analysis were carried out in order to develop production forecasts it accomplished by making numerous numbers of models based on various technology. It is inefficient in predicting few technologies moreover Wind. The first step in future work will investigate the demeanour of the neural network. One of the appealing future works will be analyzing the key execution gauge of diverse technology by creating numerous more models with various parameters within the network in conjunction with miscellaneous meteorological data testing them at numerous times in the neural networks. This leads to including study execution afterwards organize to develop one standard model for each technology with minimum error defect.

10. References

- [1] (19/12/2019). *7th environmental action programme* Available: <https://ec.europa.eu/environment/action-programme/>
- [2] *Portuguese of renewable energy* Available: <https://www.apren.pt/pt/energias-renovaveis/o-que-sao/>
- [3] *European cross-border intraday market.* Available: <http://www.mercado.ren.pt/EN/Electr/InterProj/XBID/Pages/default.aspx>
- [4] *Entidade reguladora dos servicos energeticos [erse] -the mibel liberalization of the electric market.* Available: <https://www.erse.pt/en/electricity/the-sector/>
- [5] M. R. Council, "Description of the operation of the mibel," *Retrieved March*, vol. 19, p. 2014, 2009.
- [6] E. Union, "Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC," *Official Journal of the European Union*, vol. 5, p. 2009, 2009.
- [7] N. P. Spot and N. P. OMEL, "Price Coupling of Regions (PCR)," ed, 2016.
- [8] J. C. P. Franco, "Armazenamento e partilha de energia," 2017.
- [9] J. L. Chaves, "Agregação de Produção Renovável no Âmbito da Produção em Regime de Mercado em Portugal," 2019.
- [10] J. F. O. Nascimento, "Estudo de preços de energia no mercado spot e futuros no MIBEL," 2017.
- [11] P. Ferreira, M. Araújo, and M. O'Kelly, "An overview of the Portuguese electricity market," *Energy Policy*, vol. 35, no. 3, pp. 1967-1977, 2007.
- [12] S. M. C. Conduto, "Rentabilidade e risco na comercialização agregada de energia elétrica," Instituto Superior de Engenharia de Lisboa, 2013.
- [13] (14-02-2020). *MARKET INFORMATION SYSTEM REN-Redes Energéticas Nacionais.* Available: <https://www.ren.pt/en-GB/o-que-fazemos/eletricidade/sistema-de-informacao-de-mercados>
- [14] (15/02/2020). *Estiablishment of MIBEL- Iberian Electricity Market.* Available: https://www.mibel.com/en/home_en/
- [15] 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.
- [16] *OMIE is the nominated electricity market operator.* Available: <https://www.omie.es/en/sobre-nosotros>
- [17] J. R. Andrade, J. Filipe, M. Reis, and R. J. Bessa, "Probabilistic price forecasting for day-ahead and intraday markets: Beyond the statistical model," *Sustainability*, vol. 9, no. 11, p. 1990, 2017.
- [18] E. SPOT, "Cross-border intraday: Questions & answers," *Paris: EPEX SPOT SE.* <https://www.epexspot.com/document/40068/XBID%20Q%26A>. *Zugegriffen am*, vol. 2, p. 2019, 2018.
- [19] (12-01-2020). *Overall System Management (GGS).* Available: <http://www.mercado.ren.pt/EN/Electr/MarketInfo/MarketResults/ReserveEnergy/Pages/default.aspx>
- [20] M. Bueno-Lorenzo, M. Á. Moreno, and J. Usaola, "Analysis of the imbalance price scheme in the Spanish electricity market: A wind power test case," *Energy policy*, vol. 62, pp. 1010-1019, 2013.
- [21] C. Ensslin, *The Influence of Modelling Accuracy on the Determination of Wind Power Capacity Effects and Balancing Needs.* kassel university press GmbH, 2006.

- [22] Germany energy market imbalance. Available: <https://www.regelleistung.net/ext/static/rebap?lang=en>
- [23] M. Madrigal and K. Porter, *Operating and planning electricity grids with variable renewable generation: Review of emerging lessons from selected operational experiences and desktop studies*. The World Bank, 2013.
- [24] Spanish imbalance electricity market Available: <https://www.ree.es/en/activities/operation-of-the-electricity-system>
- [25] (20-01-2020). Negative imbalance price. Available: <https://energyanalyst.co.uk/negative-prices-in-european-power-markets/>
- [26] A. T. S. Aurélie Faure (IFRI), Dražen Jakšić (EIHP), "Electricity Security of supply and Capacity Remuneration Schemes," September 2016.
- [27] (6/3/2020). Renewable-energy-advantages-and-disadvantages. Available: <https://news.energysage.com/advantages-and-disadvantages-of-renewable-energy/>
- [28] (05-01-2020). Time series Available: <https://towardsdatascience.com/an-overview-of-time-series-forecasting-models-a2fa7a358fcb>
- [29] C. Chatfield, *Time-series forecasting*. Chapman and Hall/CRC, 2000.
- [30] (Energy forecasting). Available: https://en.wikipedia.org/wiki/Energy_forecasting
- [31] J. P. M. Conde, "Previsão de preços de eletricidade no mercado diário e intradiário mibel," 2015.
- [32] "Forecasting Terms and techniques."
- [33] S. L. Vandure, M. Ramannavar, and N. S. Sidnal, "Trend Projection using Predictive Analytics," *International Journal of Computer Applications*, vol. 975, p. 8887, 2014.
- [34] (20/04/2020). Trend Projection Forecasting Methods Available: <https://businessjargons.com/trend-projection-method.html>
- [35] C. Molnar, *Interpretable machine learning*. Lulu. com, 2019.
- [36] (25/03/2020). Interpretable machine learning web page. Available: <https://christophm.github.io/interpretable-ml-book/limo.html>
- [37] G. P. Zhang, *Neural networks in business forecasting*. IGI global, 2004.
- [38] É. H. des Travaux and P. Km, "Causal Method and Time Series Forecasting model based on Artificial Neural Network," *International Journal of Computer Applications*, vol. 975, p. 8887.
- [39] J. P. d. S. Catalão, S. J. P. S. Mariano, V. Mendes, and L. Ferreira, "Short-term electricity prices forecasting in a competitive market: A neural network approach," *electric power systems research*, vol. 77, no. 10, pp. 1297-1304, 2007.
- [40] F. Rodrigues, C. Cardeira, and J. M. F. Calado, "The daily and hourly energy consumption and load forecasting using artificial neural network method: a case study using a set of 93 households in Portugal," *Energy Procedia*, vol. 62, pp. 220-229, 2014.
- [41] P. Wallström, "Evaluation of forecasting techniques and forecast errors: with focus on intermittent demand," Luleå tekniska universitet, 2009.
- [42] (20/12/2019-20/09/2020). Kisense Energy Optimiser Available: <https://kisense.s-energy.pt/?redirect=%252F>
- [43] (24-12-2019/20-09-2020). Meteorology (weather Data). Available: <http://www.soda-pro.org/web-services/meteo-data/merra>
- [44] P. g. i. system. (12-12-2019/20-09-2020). Available: https://re.jrc.ec.europa.eu/pvg_tools/en/#MR
- [45] Pearson's correlation coefficient limit. Available: <https://www.statisticssolutions.com/pearsons-correlation-coefficient/>
- [46] (1/1/2020). Pearson's correlation coefficient formula. Available: <https://www.statisticshowto.com/probability-and-statistics/correlation-coefficient-formula/#Pearson>
- [47] (02/02/2020). Pearson correlation coefficient. Available: https://en.wikipedia.org/wiki/Pearson_correlation_coefficient#Definition

- [48] K. Suzuki, *Artificial neural networks: methodological advances and biomedical applications*. BoD—Books on Demand, 2011.
- [49] A. Hajian and P. Styles, *Application of soft computing and intelligent methods in geophysics*. Springer, 2018.
- [50] L. M. Moreira, "Previsão probabilística de preços de eletricidade para o mercado diário MIBEL," 2016.

ANNEX 1

Wind Technology

Weather data Specification

Modern-Era Retrospective Analysis for Research and Applications (MERRA)-version 2 Provider National Aeronautics and Space Administration (NASA) / Goddard Space Flight Center More information at- http://gmao.gsfc.nasa.gov/reanalysis/MERRA-2	
Site latitude (positive means North)	39
Site longitude (positive means East)	-9
Site altitude (m)	71
Date beginning	01-01-2018
Date end	31-12-2018
Time reference (hour)	UT
Summarization (period of integration)	Hour (h)
Date - Instant The date or instant given	for each value corresponds to the end of the summarization
Columns:	
Temperature (K)	Temperature at 2 m above ground
Relative humidity (%)	Relative humidity at 2 m above ground
Pressure (hPa)	Pressure at ground level
Wind speed (m/s)	Wind speed at 10 m above ground
Wind direction (deg)	Wind direction at 10 m above ground (0 means from North 90 from East...)
Rainfall (kg/m2)	Rainfall (= rain depth in mm)
Snowfall (kg/m2)	Snowfall
Snow depth (m)	Snow depth
Short-wave irradiation (Wh/m2)	Surface incoming shortwave irradiation (broadband)

Weather Data for Wind Technology

Date	Time	Temperature	Relative Humidity	Pressure	Wind speed (WS)	Wind direction (WD)	Pluviosity (P)	Snowfall
01-01-2018	00:00	283.95	84.62	1025.58	0.93	269.78	0	0
01-01-2018	01:00	283.32	88.77	1025.79	1.25	246.31	0	0
01-01-2018	02:00	282.41	94.95	1025.98	1.73	225.22	0.001037	0
01-01-2018	03:00	281.63	100.46	1026.04	2.27	216.05	0.007005	0
01-01-2018	04:00	281.19	103.48	1026.01	2.66	217.11	0.011974	0
01-01-2018	05:00	281.27	102.74	1026.45	2.73	228.98	0.014146	0
01-01-2018	06:00	281.46	101.13	1026.85	2.78	237.97	0.013421	0
01-01-2018	07:00	281.8	98.86	1027.08	2.86	241.49	0.012226	0
01-01-2018	08:00	282.94	92.88	1027.35	2.96	249.1	0.007619	0
01-01-2018	09:00	285.61	80.69	1028.12	3.52	261.64	0.001431	0
01-01-2018	10:00	287.62	69.52	1028.51	4.12	277.06	0.001162	0
01-01-2018	11:00	288.73	64.12	1028.11	4.36	280.77	0.001103	0
01-01-2018	12:00	289.36	61.79	1027.36	4.5	282.3	0.000878	0
01-01-2018	13:00	289.64	60.85	1026.62	4.61	284.68	0.000844	0
01-01-2018	14:00	289.53	61.43	1026.14	4.51	285.26	0.002266	0
01-01-2018	15:00	289.09	63.46	1026.03	4.25	284.46	0.002798	0
01-01-2018	16:00	287.79	70.66	1026.06	3.21	280.83	0.003467	0
01-01-2018	17:00	286.35	78.35	1026.43	2.96	280.21	0.004224	0
01-01-2018	18:00	286.33	78.59	1027.02	2.89	284.07	0.002959	0
01-01-2018	19:00	286.02	79.92	1027.52	2.82	288.63	0.001698	0
01-01-2018	20:00	285.53	81.51	1027.81	2.82	292.25	0.000772	0
01-01-2018	21:00	285.01	82.74	1027.84	2.84	292.83	0.000152	0
01-01-2018	22:00	284.22	86.47	1027.88	2.96	289.94	0.000094	0
01-01-2018	23:00	283.66	89.93	1027.91	3.07	284.61	0.000237	0

Historical Data variables

Date	Hour	Month	Actual Prod	Historical or chromology variable					
				D-1	D-2	D-3	D-4	D-5	D-6
01-01-2018	23:00	1	76462.500						
02-01-2018	23:00	1	92287.500	76462.500					
03-01-2018	23:00	1	97797.500	92287.500	76462.500				
04-01-2018	23:00	1	56945.000	97797.500	92287.500	76462.500			
05-01-2018	23:00	1	220950.000	56945.000	97797.500	92287.500	76462.500		
06-01-2018	23:00	1	229250.000	220950.000	56945.000	97797.500	92287.500	76462.500	
07-01-2018	23:00	1	381922.500	229250.000	220950.000	56945.000	97797.500	92287.500	76462.500
08-01-2018	23:00	1	112470.000	381922.500	229250.000	220950.000	56945.000	97797.500	92287.500
09-01-2018	23:00	1	163795.000	112470.000	381922.500	229250.000	220950.000	56945.000	97797.500
10-01-2018	23:00	1	331687.500	163795.000	112470.000	381922.500	229250.000	220950.000	56945.000
11-01-2018	23:00	1	264680.000	331687.500	163795.000	112470.000	381922.500	229250.000	220950.000
12-01-2018	23:00	1	41182.500	264680.000	331687.500	163795.000	112470.000	381922.500	229250.000
13-01-2018	23:00	1	215970.000	41182.500	264680.000	331687.500	163795.000	112470.000	381922.500
14-01-2018	23:00	1	125900.000	215970.000	41182.500	264680.000	331687.500	163795.000	112470.000
15-01-2018	23:00	1	36107.500	125900.000	215970.000	41182.500	264680.000	331687.500	163795.000
16-01-2018	23:00	1	127000.000	36107.500	125900.000	215970.000	41182.500	264680.000	331687.500

Correlation of Historical vs Production using Pearson Coefficient

Production	D-1	Production	D-2	Production 2	D-3	Correlation Value
92287.5000	76462.5000	97797.5000	76462.5000	56945.0000	76462.5000		Production Vs D-1
97797.5000	92287.5000	56945.0000	92287.5000	220950.0000	92287.5000		0.4785
56945.0000	97797.5000	220950.0000	97797.5000	229250.0000	97797.5000		Production Vs D-2
220950.0000	56945.0000	229250.0000	56945.0000	381922.5000	56945.0000		0.2452
229250.0000	220950.0000	381922.5000	220950.0000	112470.0000	220950.0000		Production Vs D-3
381922.5000	229250.0000	112470.0000	229250.0000	163795.0000	229250.0000		0.1344
112470.0000	381922.5000	163795.0000	381922.5000	331687.5000	381922.5000		
163795.0000	112470.0000	331687.5000	112470.0000	264680.0000	112470.0000		
331687.5000	163795.0000	264680.0000	163795.0000	41182.5000	163795.0000		

Meteorological Data

Date	Hour	Month	Actual Prod	Metrological variable					
				Avg WS (D)	Min WS (D)	Max of WS	Avg WD (D)	Min WD (D)	Max WD (D)
01-01-2018	23:00	1	76462.500	3.07	0.93	4.61	265.90	216.05	292.83
02-01-2018	23:00	1	92287.500	3.33	2.41	4.54	269.00	258.08	279.8
03-01-2018	23:00	1	97797.500	3.10	2.34	3.94	302.29	274.12	326.55
04-01-2018	23:00	1	56945.000	3.52	2.2	5.13	268.73	227.24	311.29
05-01-2018	23:00	1	220950.000	6.15	4.62	8.38	283.47	215.37	342.1
06-01-2018	23:00	1	229250.000	5.07	3.38	7.42	341.02	332.43	353.72
07-01-2018	23:00	1	381922.500	6.50	4.62	9.62	345.20	336.52	354.16
08-01-2018	23:00	1	112470.000	3.25	0.66	5.01	152.54	5.03	350.34
09-01-2018	23:00	1	163795.000	5.10	3.51	8.28	227.27	180.18	308.76
10-01-2018	23:00	1	331687.500	5.51	3.75	9.75	279.03	237.55	318.8
11-01-2018	23:00	1	264680.000	4.70	1.78	6.79	291.51	2.64	357.98
12-01-2018	23:00	1	41182.500	3.55	1.22	6.3	145.55	50.64	182.56
13-01-2018	23:00	1	215970.000	5.08	2.77	6.72	268.81	161.59	317.99
14-01-2018	23:00	1	125900.000	3.62	2.83	4.9	264.66	5.66	350.42
15-01-2018	23:00	1	36107.500	2.68	1.31	3.33	167.60	7.87	356.79
16-01-2018	23:00	1	127000.000	2.91	1.38	3.95	305.82	271.01	353.8
17-01-2018	23:00	1	308672.500	5.38	4.08	7.25	138.61	0.2	359.59

Correlation of Meteorological vs Production using Pearson Coefficient

Metrological variable						Correlation Value
Production	Avg WS (D)	Production	Min WS (D)	Production 2	Max of WS (D)	
76462.5000	3.07	76462.5000	0.93	76462.5000	4.61	Production Vs Avg-Wind Speed (WS)
92287.5000	3.33	92287.5000	2.41	92287.5000	4.54	0.77124902
97797.5000	3.10	97797.5000	2.34	97797.5000	3.94	Production Vs Min-Wind Speed (WS)
56945.0000	3.52	56945.0000	2.20	56945.0000	5.13	0.749641731
220950.0000	6.15	220950.0000	4.62	220950.0000	8.38	Production Vs Max-Wind Speed (WS)
229250.0000	5.07	229250.0000	3.38	229250.0000	7.42	0.636861985
381922.5000	6.50	381922.5000	4.62	381922.5000	9.62	
112470.0000	3.25	112470.0000	0.66	112470.0000	5.01	
163795.0000	5.10	163795.0000	3.51	163795.0000	8.28	
331687.5000	5.51	331687.5000	3.75	331687.5000	9.75	
264680.0000	4.70	264680.0000	1.78	264680.0000	6.79	
41182.5000	3.55	41182.5000	1.22	41182.5000	6.30	
215970.0000	5.08	215970.0000	2.77	215970.0000	6.72	
125900.0000	3.62	125900.0000	2.83	125900.0000	4.90	
36107.5000	2.68	36107.5000	1.31	36107.5000	3.33	
127000.0000	2.91	127000.0000	1.38	127000.0000	3.95	

Best Model Result of Wind Technology from Hourly and Daily Models

MPD 03			Target	Input/Sample		Training	Results
Date	Hour	Month	Actual Production (A)	D-1	Avg WS (D)	Simulation	MAPE
02-01-2018	23:00	1	92287.500	76462.500	3.33	Simulation 01	56.6666
03-01-2018	23:00	1	97797.500	92287.500	3.10		
04-01-2018	23:00	1	56945.000	97797.500	3.52	Simulation 02	62.51543
05-01-2018	23:00	1	220950.000	56945.000	6.15		
06-01-2018	23:00	1	229250.000	220950.000	5.07	Simulation 03	51.57913
14-01-2018	23:00	1	125900.000	215970.000	3.62		
15-01-2018	23:00	1	36107.500	125900.000	2.68	Simulation 04	56.7615
16-01-2018	23:00	1	127000.000	36107.500	2.91		
17-01-2018	23:00	1	308672.500	127000.000	5.38	Simulation 05	55.35624
18-01-2018	23:00	1	209807.500	308672.500	5.07		
19-01-2018	23:00	1	395527.500	209807.500	5.92	Avg Simulation	55.31016

ANNEX 2

Solar Technology

Weather data Specification

Modern-Era Retrospective Analysis for Research and Applications (MERRA)-version 2 Provider National Aeronautics and Space Administration (NASA) / Goddard Space Flight Center More information at http://gmao.gsfc.nasa.gov/reanalysis/MERRA-2	
Site latitude (positive means North)	37.2
Site longitude (positive means East)	-7.9
Site altitude (m)	54
Date beginning	01-01-2018
Date end	31-12-2018
Time reference (hour)	UT
Summarization (period of integration)	Hour (h)
Date - Instant The date or instant given	for each value corresponds to the end of the summarization
Columns:	
Temperature (K)	Temperature at 2 m above ground
Relative humidity (%)	Relative humidity at 2 m above ground
Pressure (hPa)	Pressure at ground level
Wind speed (m/s)	Wind speed at 10 m above ground
Wind direction (deg)	Wind direction at 10 m above ground (0 means from North 90 from East...)
Rainfall (kg/m2)	Rainfall (= rain depth in mm)
Snowfall (kg/m2)	Snowfall
Snow depth (m)	Snow depth
Short-wave irradiation (Wh/m2)	Surface incoming shortwave irradiation (broadband)

Weather Data for Solar Technology

# Date	UT tin	Temperature (C)	Relative Humidity	Pressure	Wind speed	Wind direction	Pluviosity (P)	Snowfall
01-01-2018	00:00	282.27	97.97	1020.59	3.57	12.8	0.000795	0
01-01-2018	01:00	282	97.76	1020.64	2.99	20.76	0.001984	0
01-01-2018	02:00	281.96	96.54	1021	2.67	25.91	0.00196	0
01-01-2018	03:00	282.03	94.77	1021.21	2.37	30.29	0.002923	0
01-01-2018	04:00	282.12	92.63	1021.02	2.27	34.93	0.006083	0
01-01-2018	05:00	281.95	92.03	1020.99	2.07	35.97	0.005596	0
01-01-2018	06:00	281.76	91.47	1021.51	2.24	28.45	0.005468	0
01-01-2018	07:00	281.66	88.95	1022.14	2.53	21.8	0.005199	0
01-01-2018	08:00	281.83	87.37	1022.65	2.62	20.07	0.003588	0
01-01-2018	09:00	282.74	82.67	1023.08	2.4	21.68	0.001215	0
01-01-2018	10:00	284.29	74.61	1023.48	2.62	11.52	0.000131	0
01-01-2018	11:00	286.35	63.89	1023.51	3.49	3.87	0	0
01-01-2018	12:00	288.02	53.7	1022.53	3.55	5.67	0	0
01-01-2018	13:00	288.97	47.87	1021.43	3.05	6.63	0	0
01-01-2018	14:00	289.47	46.01	1020.85	2.97	359.85	0	0
01-01-2018	15:00	289.41	47.91	1020.74	3.13	353.5	0	0
01-01-2018	16:00	288.06	59.89	1020.66	2.71	351.7	0	0
01-01-2018	17:00	285.89	67.72	1020.77	3.2	350.32	0	0
01-01-2018	18:00	285.08	73.74	1021.27	3.66	350.3	0.000001	0
01-01-2018	19:00	284.47	80.13	1021.92	4.1	351.35	0.000002	0
01-01-2018	20:00	283.87	86.82	1022.42	4.37	355.31	0.000007	0
01-01-2018	21:00	283.28	92.96	1022.71	4.48	2.01	0.000017	0

Solar Radiation Data

						Latitude (decimal degrees):37.200
						Longitude (decimal degrees):-7.900
						Radiation database:PVGIS-SARAH
						Optimal slope angle (deg.):32
No.	Month	Avg-Hh	Avg-Hi_op	Avg-Hi_m	Avg-Hb(n)_m	
1	Jan 2005-2016	80.4975	132.4925	123.1475	131.9583	
2	Feb 2005-2016	95.27917	134.4075	128.0725	122.8125	
3	Mar 2005-2016	141.9117	172.1158	168.6508	153.9258	
4	Apr 2005-2016	173.1358	182.7258	184.2983	168.6200	
5	May 2005-2016	215.7983	206.6725	212.9808	213.9058	
6	Jun 2005-2016	231.235	210.8075	219.7183	236.5417	
7	Jul 2005-2016	250.2183	233.3008	242.2575	279.4283	
8	Aug 2005-2016	221.1717	227.0742	230.8508	249.0867	
9	Sep 2005-2016	163.9375	189.94	187.98	184.8200	
10	Oct 2005-2016	118.3892	157.9917	152.0542	144.3675	
11	Nov 2005-2016	83.5375	130.0333	121.895	122.1167	
12	Dec 2005-2016	71.285	121.1475	112.035	117.8050	

Historical Data variables

Date	Hour	Month	Actual Production	Historical or chromology variable					
				D-1	D-2	D-3	D-4	D-5	D-6
01-01-2018	23:00	1	9372.000						
02-01-2018	23:00	1	6779.000	9372.000					
03-01-2018	23:00	1	2713.250	6779.000	9372.000				
04-01-2018	23:00	1	2992.750	2713.250	6779.000	9372.000			
05-01-2018	23:00	1	939.250	2992.750	2713.250	6779.000	9372.000		
06-01-2018	23:00	1	12498.250	939.250	2992.750	2713.250	6779.000	9372.000	
07-01-2018	23:00	1	7272.500	12498.250	939.250	2992.750	2713.250	6779.000	9372.000
08-01-2018	23:00	1	11039.500	7272.500	12498.250	939.250	2992.750	2713.250	6779.000
09-01-2018	23:00	1	818.500	11039.500	7272.500	12498.250	939.250	2992.750	2713.250
10-01-2018	23:00	1	4168.750	818.500	11039.500	7272.500	12498.250	939.250	2992.750
11-01-2018	23:00	1	7020.750	4168.750	818.500	11039.500	7272.500	12498.250	939.250
12-01-2018	23:00	1	6769.750	7020.750	4168.750	818.500	11039.500	7272.500	12498.250
13-01-2018	23:00	1	8627.750	6769.750	7020.750	4168.750	818.500	11039.500	7272.500
14-01-2018	23:00	1	8719.750	8627.750	6769.750	7020.750	4168.750	818.500	11039.500
15-01-2018	23:00	1	12832.250	8719.750	8627.750	6769.750	7020.750	4168.750	818.500
16-01-2018	23:00	1	1686.250	12832.250	8719.750	8627.750	6769.750	7020.750	4168.750

Correlation of Historical vs Production using Pearson Coefficient

Production	D-1	Production	D-2	Production 2	D-3	Correlation Value
9372.0000	0.0000	6779.0000	0.0000	2713.2500	0.0000		Production Vs D-1
6779.0000	9372.0000	2713.2500	9372.0000	2992.7500	9372.0000		0.7084
2713.2500	6779.0000	2992.7500	6779.0000	939.2500	6779.0000		Production Vs D-2
2992.7500	2713.2500	939.2500	2713.2500	12498.2500	2713.2500		0.6639
939.2500	2992.7500	12498.2500	2992.7500	7272.5000	2992.7500		Production Vs D-3
12498.2500	939.2500	7272.5000	939.2500	11039.5000	939.2500		0.6447
7272.5000	12498.2500	11039.5000	12498.2500	818.5000	12498.2500		
11039.5000	7272.5000	818.5000	7272.5000	4168.7500	7272.5000		
818.5000	11039.5000	4168.7500	11039.5000	7020.7500	11039.5000		

Meteorological Data

Date	Hour	Actual Production (A)	Metrological variable				
			Avg T (D)	Min T (D)	Max of T (D)	Avg I (D)	Max I (D)
01-01-2018	23:00	9372.000	284.28	281.66	289.47	122.91	520.59
02-01-2018	23:00	6779.000	285.16	282.09	289.83	123.08	510.82
03-01-2018	23:00	2713.250	286.89	283.75	291.99	118.86	502.05
04-01-2018	23:00	2992.750	286.94	284.37	290.78	110.54	479.12
05-01-2018	23:00	939.250	286.65	283.23	289.68	75.42	370.61
06-01-2018	23:00	12498.250	282.85	280.91	285.97	90.63	423.61
07-01-2018	23:00	7272.500	281.98	279.98	285.36	109.43	455.02
08-01-2018	23:00	11039.500	281.67	280.29	284.84	112.21	497.69
09-01-2018	23:00	818.500	284.42	280.02	287.48	39.21	203.03
10-01-2018	23:00	4168.750	285.89	282.32	289.74	121.77	503.76
11-01-2018	23:00	7020.750	284.22	281.67	288.37	122.02	519.36
12-01-2018	23:00	6769.750	283.43	280.28	286.87	112.85	485.09
13-01-2018	23:00	8627.750	285.12	283.18	288.69	112.61	531.24
14-01-2018	23:00	8719.750	283.21	280.52	285.94	88.28	437.79
15-01-2018	23:00	12832.250	282.30	278.89	287.4	138.69	557.00
16-01-2018	23:00	1686.250	284.26	280.24	289.69	128.66	532.66

Correlation of Meteorological vs Production using Pearson Coefficient

Metrological variable						Correlation Value
Production	Avg T (D)	Production	Min T (D)	Production 2	Max of T (D)	
0.0000	284.28	0.0000	281.66	0.0000	289.47	Production Vs Avg-Temperature (T)
0.0000	284.28	0.0000	281.66	0.0000	289.47	0.207732811
0.0000	284.28	0.0000	281.66	0.0000	289.47	Production Vs Min-Temperature (T)
0.0000	284.28	0.0000	281.66	0.0000	289.47	0.178628934
0.0000	284.28	0.0000	281.66	0.0000	289.47	Production Vs Max-Temperature (T)
0.0000	284.28	0.0000	281.66	0.0000	289.47	0.224487135
0.0000	284.28	0.0000	281.66	0.0000	289.47	
0.2500	284.28	0.2500	281.66	0.2500	289.47	
296.7500	284.28	296.7500	281.66	296.7500	289.47	
1046.5000	284.28	1046.5000	281.66	1046.5000	289.47	
1508.2500	284.28	1508.2500	281.66	1508.2500	289.47	
1757.7500	284.28	1757.7500	281.66	1757.7500	289.47	
1410.2500	284.28	1410.2500	281.66	1410.2500	289.47	
1489.5000	284.28	1489.5000	281.66	1489.5000	289.47	
1203.5000	284.28	1203.5000	281.66	1203.5000	289.47	

Best Model Result of Solar Technology from Hourly and Daily Models

MPD 04			Traget	Input/Sample data					Training	Result
Date	Hour	Month	Actual Prodi	D-1	Max of T	Avg I (D)	Avg-H _n	Avg-H _{b(n)}	Simulation	MAPE
02-01-2018	23:00	1	6779.000	9372.000	289.83	123.08	80.498	131.958	Simulation 01	26.541
03-01-2018	23:00	1	2713.250	6779.000	291.99	118.86	80.498	131.958		
04-01-2018	23:00	1	2992.750	2713.250	290.78	110.54	80.498	131.958	Simulation 02	24.68251
05-01-2018	23:00	1	939.250	2992.750	289.68	75.42	80.498	131.958		
06-01-2018	23:00	1	12498.250	939.250	285.97	90.63	80.498	131.958	Simulation 03	24.36708
14-01-2018	23:00	1	8719.750	8627.750	285.94	88.28	80.498	131.958		
15-01-2018	23:00	1	12832.250	8719.750	287.4	138.69	80.498	131.958	Simulation 04	26.17638
16-01-2018	23:00	1	1686.250	12832.250	289.69	128.66	80.498	131.958		
17-01-2018	23:00	1	12818.750	1686.250	290.58	138.08	80.498	131.958	Simulation 05	25.95142
18-01-2018	23:00	1	13298.250	12818.750	290.09	138.73	80.498	131.958		
19-01-2018	23:00	1	11865.000	13298.250	289.98	138.72	80.498	131.958	Avg Simulation	23.71671
20-01-2018	23:00	1	12088.000	11865.000	290.15	142.19	80.498	131.958		
21-01-2018	23:00	1	5674.250	12088.000	290.04	139.04	80.498	131.958		
22-01-2018	23:00	1	9234.750	5674.250	289.82	137.34	80.498	131.958		
23-01-2018	23:00	1	7408.750	9234.750	290.05	147.33	80.498	131.958		
24-01-2018	23:00	1	6490.000	7408.750	288.33	145.27	80.498	131.958		
25-01-2018	23:00	1	3864.500	6490.000	287.97	88.67	80.498	131.958		
26-01-2018	23:00	1	11150.750	3864.500	287.51	150.00	80.498	131.958		

ANNEX 3

Hydro Technology

Weather data Specification

Modern-Era Retrospective Analysis for Research and Applications (MERRA)-version 2 Provider National Aeronautics and Space Administration (NASA) / Goddard Space Flight Center More information at- http://gmao.gsfc.nasa.gov/reanalysis/MERRA-2	
Site latitude (positive means North)	41.433
Site longitude (positive means East)	-7.783
Site altitude (m)	623
Date beginning	01-01-2019
Date end	31-12-2019
Time reference (hour)	UT
Summarization (period of integration)	Hour (h)
Date - Instant The date or instant given	for each value corresponds to the end of the summarization
Columns:	
Temperature (K)	Temperature at 2 m above ground
Relative humidity (%)	Relative humidity at 2 m above ground
Pressure (hPa)	Pressure at ground level
Wind speed (m/s)	Wind speed at 10 m above ground
Wind direction (deg)	Wind direction at 10 m above ground (0 means from North 90 from East...)
Rainfall (kg/m2)	Rainfall (= rain depth in mm)
Snowfall (kg/m2)	Snowfall
Snow depth (m)	Snow depth
Short-wave irradiation (Wh/m2)	Surface incoming shortwave irradiation (broadband)

Weather Data for Hydro Technology

Date	Time	Temperature	Relative Humidity	Pressure	Wind speed	Wind direction	Pluviosity (P)	Snowfall
01-01-2019	00:00	274.17	90.17	963.45	2.55	89.04	0	0
01-01-2019	01:00	273.59	90.93	963.74	2.6	86.59	0	0
01-01-2019	02:00	273.19	92.15	964.06	2.58	84.94	0	0
01-01-2019	03:00	272.95	93.42	964.03	2.54	84.03	0	0
01-01-2019	04:00	272.9	94.34	963.75	2.59	83.85	0	0
01-01-2019	05:00	272.81	95.77	963.61	2.61	84.14	0	0
01-01-2019	06:00	272.67	97.39	963.78	2.58	85.24	0	0
01-01-2019	07:00	272.55	98.54	964.09	2.51	87.01	0.000002	0
01-01-2019	08:00	273.57	93.72	964.51	2.34	88.66	0.00002	0.000039
01-01-2019	09:00	276.28	85.03	964.96	2.3	89.75	0.000031	0.000033
01-01-2019	10:00	278.84	74.32	965.11	2.03	89.63	0	0
01-01-2019	11:00	281.15	67.68	964.8	1.7	86.17	0	0
01-01-2019	12:00	283.2	62.69	964.17	1.37	76.24	0	0
01-01-2019	13:00	284.61	59.65	963.69	1.2	60.53	0	0
01-01-2019	14:00	285.21	60.43	963.58	1.1	49.8	0	0
01-01-2019	15:00	284.82	70.16	963.63	0.77	40.33	0	0
01-01-2019	16:00	283.11	70.14	963.79	0.7	17.58	0	0
01-01-2019	17:00	281.46	70.41	964.12	0.86	355.91	0	0
01-01-2019	18:00	280.63	73.38	964.5	1.12	353.28	0	0
01-01-2019	19:00	279.85	76.25	964.8	1.35	359.05	0	0
01-01-2019	20:00	279.12	78.91	964.9	1.48	5.56	0	0
01-01-2019	21:00	279.12	77.79	964.87	1.49	10.51	0	0

Historical Data variables

Date	hour	Month	Historical or chromology variable						
			Actual Production (A)	D-1	D-2	D-3	D-4	D-5	D-6
01-01-2019	23:00	1	63242.5						
02-01-2019	23:00	1	61172.5	63242.5					
03-01-2019	23:00	1	55137.5	61172.5	63242.5				
04-01-2019	23:00	1	51195	55137.5	61172.5	63242.5			
05-01-2019	23:00	1	46967.5	51195	55137.5	61172.5	63242.5		
06-01-2019	23:00	1	44862.5	46967.5	51195	55137.5	61172.5	63242.5	
07-01-2019	23:00	1	44392.5	44862.5	46967.5	51195	55137.5	61172.5	63242.5
08-01-2019	23:00	1	42957.5	44392.5	44862.5	46967.5	51195	55137.5	61172.5
09-01-2019	23:00	1	41392.5	42957.5	44392.5	44862.5	46967.5	51195	55137.5
10-01-2019	23:00	1	40107.5	41392.5	42957.5	44392.5	44862.5	46967.5	51195
11-01-2019	23:00	1	38310	40107.5	41392.5	42957.5	44392.5	44862.5	46967.5
12-01-2019	23:00	1	37060	38310	40107.5	41392.5	42957.5	44392.5	44862.5
13-01-2019	23:00	1	35237.5	37060	38310	40107.5	41392.5	42957.5	44392.5
14-01-2019	23:00	1	35052.5	35237.5	37060	38310	40107.5	41392.5	42957.5
15-01-2019	23:00	1	31127.5	35052.5	35237.5	37060	38310	40107.5	41392.5
16-01-2019	23:00	1	32080	31127.5	35052.5	35237.5	37060	38310	40107.5

Correlation of Historical vs Production using Pearson Coefficient

Production	D-1	Production	D-2	Production 2	D-3	Correlation Value
63242.5000	0.0000	61172.5000	0.0000	55137.5000	0.0000		Production Vs D-1
61172.5000	63242.5000	55137.5000	63242.5000	51195.0000	63242.5000		0.9195
55137.5000	61172.5000	51195.0000	61172.5000	46967.5000	61172.5000		Production Vs D-2
51195.0000	55137.5000	46967.5000	55137.5000	44862.5000	55137.5000		0.8295
46967.5000	51195.0000	44862.5000	51195.0000	44392.5000	51195.0000		Production Vs D-3
44862.5000	46967.5000	44392.5000	46967.5000	42957.5000	46967.5000		0.7373
44392.5000	44862.5000	42957.5000	44862.5000	41392.5000	44862.5000		
42957.5000	44392.5000	41392.5000	44392.5000	40107.5000	44392.5000		
41392.5000	42957.5000	40107.5000	42957.5000	38310.0000	42957.5000		
40107.5000	41392.5000	38310.0000	41392.5000	37060.0000	41392.5000		
38310.0000	40107.5000	37060.0000	40107.5000	35237.5000	40107.5000		
37060.0000	38310.0000	35237.5000	38310.0000	35052.5000	38310.0000		
35237.5000	37060.0000	35052.5000	37060.0000	31127.5000	37060.0000		
35052.5000	35237.5000	31127.5000	35237.5000	32080.0000	35237.5000		
31127.5000	35052.5000	32080.0000	35052.5000	31372.5000	35052.5000		
32080.0000	31127.5000	31372.5000	31127.5000	31145.0000	31127.5000		
31372.5000	32080.0000	31145.0000	32080.0000	48945.0000	32080.0000		
31145.0000	31372.5000	48945.0000	31372.5000	132577.5000	31372.5000		

Meteorological Data

Date	hour	Month	Actual Production	Metrological variable					
				sum P (D)	Min P (D)	Max of P (D)	Avg H (D)	Min H (D)	Max H (D)
01-01-2019	23:00	1	63242.5	0.000053	0	0.000031	80.272	59.65	98.54
02-01-2019	23:00	1	61172.5	0.000431	0	0.000203	86.542	61.59	109.83
03-01-2019	23:00	1	55137.5	0.000032	0	0.000004	79.739	54.96	96.99
04-01-2019	23:00	1	51195	0	0	0	74.857	55.25	84.51
05-01-2019	23:00	1	46967.5	0	0	0	66.708	48.71	80.11
06-01-2019	23:00	1	44862.5	0	0	0	71.181	50.54	85.79
07-01-2019	23:00	1	44392.5	0.000001	0	0.000001	74.518	56.2	85.72
08-01-2019	23:00	1	42957.5	0.000061	0	0.000009	80.516	56.89	96.73
09-01-2019	23:00	1	41392.5	0.000935	0	0.000091	88.967	66.28	102.18
10-01-2019	23:00	1	40107.5	0.00125	0.000001	0.000316	83.710	58.2	101.04
11-01-2019	23:00	1	38310	0.000366	0	0.000045	83.622	60.77	100.46
12-01-2019	23:00	1	37060	0.001849	0.000008	0.000641	89.430	63.78	103.09
13-01-2019	23:00	1	35237.5	0.001256	0	0.000732	89.228	63.94	105.43
14-01-2019	23:00	1	35052.5	0.008934	0	0.006554	91.167	66.6	110.09
15-01-2019	23:00	1	31127.5	0.014456	0	0.003206	86.800	68.91	96.3
16-01-2019	23:00	1	32080	2.042394	0.00062	0.169913	98.060	94.13	102.43
17-01-2019	23:00	1	31372.5	0.337189	0.000166	0.087042	88.474	55.84	108.31
18-01-2019	23:00	1	31145	3.2689	0.000219	0.499732	95.627	88.13	105.72
19-01-2019	23:00	1	48945	16.738313	0.029181	1.908076	96.675	94.66	98.26
20-01-2019	23:00	1	132577.5	0.540681	0.008796	0.074361	92.415	72.55	103.58

Correlation of Meteorological vs Production using Pearson Coefficient

Metrological variable						Correlation Value
Production	Sum P	Production	Min P	Production 2	Max of P
63242.5000	0.0001	63242.5000	0.0000	63242.5000	0.0000	Production Vs Avg-Temperature (T)
61172.5000	0.0004	61172.5000	0.0000	61172.5000	0.0002	0.449957844
55137.5000	0.0000	55137.5000	0.0000	55137.5000	0.0000	Production Vs Min-Temperature (T)
51195.0000	0.0000	51195.0000	0.0000	51195.0000	0.0000	0.455795464
46967.5000	0.0000	46967.5000	0.0000	46967.5000	0.0000	Production Vs Max-Temperature (T)
44862.5000	0.0000	44862.5000	0.0000	44862.5000	0.0000	0.32606132
44392.5000	0.0000	44392.5000	0.0000	44392.5000	0.0000	
42957.5000	0.0001	42957.5000	0.0000	42957.5000	0.0000	
41392.5000	0.0009	41392.5000	0.0000	41392.5000	0.0001	
40107.5000	0.0013	40107.5000	0.0000	40107.5000	0.0003	
38310.0000	0.0004	38310.0000	0.0000	38310.0000	0.0000	
37060.0000	0.0018	37060.0000	0.0000	37060.0000	0.0006	
35237.5000	0.0013	35237.5000	0.0000	35237.5000	0.0007	
35052.5000	0.0089	35052.5000	0.0000	35052.5000	0.0066	
31127.5000	0.0145	31127.5000	0.0000	31127.5000	0.0032	
32080.0000	2.0424	32080.0000	0.0006	32080.0000	0.1699	
31372.5000	0.3372	31372.5000	0.0002	31372.5000	0.0870	
31145.0000	3.2689	31145.0000	0.0002	31145.0000	0.4997	

Best Model Result of Hydro Technology from Hourly and Daily Models

MPD 03	Traget			Input/Sample data					Training	Result
Date	hour	Month	Actual Prodi	D-1	Min P (D-1)	Min H (D)	Sum P (D-6)	Avg H (D-6)	Simulation	MAPE
13-01-2019	23:00	1	35237.5	37060	0.000008	63.78	0.0045	83.460	Simulation 01	25.67707
14-01-2019	23:00	1	35052.5	35237.5	0	63.94	0.0057	85.912		
15-01-2019	23:00	1	31127.5	35052.5	0	66.6	0.0146	87.687	Simulation 02	21.97311
16-01-2019	23:00	1	32080	31127.5	0	68.91	0.0281	87.326		
17-01-2019	23:00	1	31372.5	32080	0.00062	94.13	2.0693	89.718	Simulation 03	26.79493
18-01-2019	23:00	1	31145	31372.5	0.000166	55.84	2.4061	90.526		
19-01-2019	23:00	1	48945	31145	0.000219	88.13	5.6731	91.559	Simulation 04	24.79202
20-01-2019	23:00	1	132577.5	48945	0.029181	94.66	22.41	92.800		
21-01-2019	23:00	1	57907.5	132577.5	0.008796	72.55	22.942	93.008	Simulation 05	26.32572
22-01-2019	23:00	1	52905	57907.5	0.000305	58.45	23.029	93.100		
23-01-2019	23:00	1	54067.5	52905	0.000707	93.16	23.568	92.802	Avg Simulation	21.43701
24-01-2019	23:00	1	50710	54067.5	0.049269	82.27	26.26	93.571		
25-01-2019	23:00	1	45850	50710	0.006276	76.1	23.789	93.142		
26-01-2019	23:00	1	43725	45850	0	58.21	7.1226	91.326		
27-01-2019	23:00	1	43202.5	43725	0	59.68	6.7124	90.935		
28-01-2019	23:00	1	42682.5	43202.5	0.000855	62.88	7.2294	91.055		
29-01-2019	23:00	1	87152.5	42682.5	0.001647	82.64	6.7327	90.600		

Data Variable Abbreviation

Temperature (T)	Avg Temperature (AvgT)
	Min Temperature (MinT)
	Max Temperature (MaxT)
Wind speed (WS)	Avg Wind speed (AvgWS)
	Min Wind speed (MinWS)
	Max Wind speed (MaxWS)
Wind direction (WD)	Avg Wind direction (AvgWD)
	Min Wind direction (MinWD)
	Max Wind direction (MaxWD)
Rainfall Snowfall/Pluviosity (P)	Avg Pluviosity (AvgP)
	Min Pluviosity (MinP)
	Max Pluviosity (MaxP)
Short-wave irradiation/ irradiation(I)	Avg irradiation(Avg I)
	Min irradiation(Min I)
	Max irradiation(Max I)
Relative Humidity(H)	Avg Humidity(AvgH)
	Min Humidity(MinH)
	Max Humidity(MaxH)