

Energy Evaluation And Improvement Of Silk House Microgrid System

Mohamed Elrhoul

Work carried out in the Double Degree program between the University of A Coruña and the Polytechnic Institute of Bragança, as a requirement for obtaining the Master's Degree respectively in Energy Efficiency and Utilization and Renewable Energies and Energy Efficiency.

Supervisors

Prof. Ana Isabel Pereira

Prof. José Luís Calvo Rolle

Prof. Esteban Jove Pérez

Bragança

2022

Acknowledgements

I would like to thank the following people who have helped me undertake this research:

*My supervisor Professor **Ana Isabel Pereira**, for his enthusiasm for the project, for his support, encouragement and patience.*

*I would also like to express my gratitude to Professor **José Luis Calvo Rolle**, for his judicious advice.*

*I would also like to thank the Phd student Mr. **Yahia Amoura**, for his encouragement and support.*

*Many thanks to academic members of **University of Coruña** and **Instituto Politécnico de Bragança**.*

I would like to express my gratitude to the members of the juries that agreed to evaluate my work.

Dedication

To my dearest mother,

Whatever I do or say to you, I cannot thank you as I should, your affection covers me, your benevolence guides me and your presence close to me has always been my source of strength to face the different obstacles.

To my dearest father,

You have always been close to me to support and encourage me. May this work reflect my gratitude and affection.

To my dearest sister Doha,

For your presence, your love, your sacrifices and your encouragement throughout my university career. Thank you.

To my dearest brother Abdessamie,

May the god give you health, happiness, courage and especially success.

Abstract

In solar-powered microgrids, solar power yield prediction is critical to optimize operations and reduce the impact of uncertainties. Solar radiance prediction is essential for solar power yield prediction, which typically requires historical solar radiance data. Moreover, uncertainty and low correlation of Photovoltaic power availability with load reduces the benefits of their integration in the microgrid. These challenges can be resolved by implementing buffer such as an energy storage system or a diesel group. However this leads to increased operational cost. Solar radiance forecasting helps to reduce reserve requirement, thereby improving the use of photovoltaic energy inside the microgrid system, technically in term of power control, economically in term of energy cost and environmentally in term of toxic emissions that can be provided from the buffer systems such as the batteries and diesel group. The work presented in this thesis addresses the problem of predicting solar radiance in order to increase the performance of the Bragança city silk house microgrid using regression machine learning based approaches including linear regression models, regression trees, Gaussian process regression models, support vector machines, kernel approximation, ensembles of regression trees, and neural network regression models, the model performances results are compared in term root mean squared errors, mean square error, R-squared, mean absolute error.

Keywords:

Renewable Energies, Microgrid, Solar Irradiation, Forecasting, Data Correlation, Regression Models, Root Mean Squared Error.

Resumo

Nas micro redes de energia solar, a previsão da produção de energia solar é fundamental para otimizar as operações e reduzir o impacto das incertezas. A previsão da irradiação solar é essencial para a previsão da produção de energia solar, o que normalmente requer dados históricos de irradiação solar. Além disso, a incerteza e a baixa correlação da disponibilidade de energia fotovoltaica com a carga reduz os benefícios da sua integração no microgrid. Estes desafios podem ser resolvidos através da implementação de amortecedores, tais como um sistema de armazenamento de energia ou um grupo diesel. No entanto, isto leva a um aumento do custo operacional. A previsão da irradiação solar ajuda a reduzir a necessidade de reservas, melhorando assim a utilização de energia fotovoltaica dentro do sistema de microgeração, tecnicamente em termos de controlo de energia, economicamente em termos de custo energético e ambientalmente em termos de emissões tóxicas que podem ser fornecidas pelos sistemas tampão, tais como as baterias e o grupo diesel. O trabalho apresentado nesta tese aborda o problema da previsão da irradiação solar a fim de aumentar o desempenho da micro-rede da casa da seda da cidade de Bragança utilizando abordagens baseadas na aprendizagem da máquina de regressão, incluindo modelos de regressão linear, árvores de regressão, modelos de regressão do processo Gaussiano, máquinas vectoriais de suporte, aproximação do núcleo, conjuntos de árvores de regressão, e modelos de regressão da rede neural, os resultados do desempenho dos modelos são comparados em termos de erros quadráticos médios de raiz, erro quadrático médio, R-quadrado, erro absoluto médio..

Palavras-chave:

Energias renováveis, Micro-rede, Irradiação solar, Previsão, Correlação de Dados, Modelos de Regressão, Erro de Raiz Média Quadrática.

Resumen

En las microrredes alimentadas por energía solar, la predicción del rendimiento de la energía solar es fundamental para optimizar las operaciones y reducir el impacto de las incertidumbres. La predicción de la irradiación solar es esencial para la predicción del rendimiento de la energía solar, que normalmente requiere datos históricos de irradiación solar. Además, la incertidumbre y la baja correlación de la disponibilidad de la energía fotovoltaica con la carga reducen los beneficios de su integración en el microgrupo. Estos problemas pueden resolverse implementando un sistema de almacenamiento de energía o un grupo diesel. Sin embargo, esto conlleva un aumento de los costes operativos. La previsión de la irradiación solar ayuda a reducir las necesidades de reserva, mejorando así el uso de la energía fotovoltaica dentro del sistema de la microrred, técnicamente en términos de control de potencia, económicamente en términos de coste energético y medioambientalmente en términos de emisiones tóxicas que pueden proporcionar los sistemas de amortiguación como las baterías y el grupo diesel. El trabajo presentado en esta tesis aborda el problema de la predicción de la irradiación solar con el fin de aumentar el rendimiento de la microrred de la casa de seda de la ciudad de Braganza, utilizando enfoques basados en el aprendizaje automático de regresión, incluyendo modelos de regresión lineal, árboles de regresión, modelos de regresión de procesos gaussianos, máquinas de vectores de apoyo, aproximación de kernel, conjuntos de árboles de regresión y modelos de regresión de redes neuronales, error cuadrático medio, R-cuadrado, error absoluto medio.

palabras-clave:

Energías renovables, microrredes, irradiación solar, previsión, correlación de datos, modelos de regresión, error medio cuadrático.

Contents

1	General Introduction	1
1.1	Introduction	1
1.2	Motivations	2
1.3	Literature Review	3
1.4	Objectives	3
1.5	Thesis Structure	4
2	State of The Art	5
2.1	Introduction	5
2.2	Renewable Energy	5
2.3	Renewable Energy Technologies	6
2.3.1	Photovoltaic Power	6
2.3.2	Wind Energy	7
2.3.3	Hydroelectric Power	7
2.4	Microgrid Management System	8
2.5	Machine Learning	9
2.6	Errors	11
2.6.1	Mean Square Error	11
2.6.2	Root Mean Square Error	12
2.6.3	R-squared	12
2.6.4	Mean Absolute Error	13
3	Silk house Building	14
3.1	Introduction	14
3.2	Electrical Characterization of the Building	15
3.3	Characterization of the Energy Potential of the Building	15
3.4	Microgrid in Silk house	16
3.4.1	Photovoltaic system	17

3.4.2	Pico-Hydro power	17
4	Case of Study	19
4.1	Introduction	19
4.2	Data Characterization	20
4.3	Data Correlation	22
5	Results and Discussions	26
5.1	Introduction	26
5.2	Models Training and Validation	27
5.2.1	Without Data Treatment	27
5.2.2	With Data Treatment	28
5.3	GPR model Exporting and Testing	34
6	General Conclusion and Perspectives	36
	Bibliography	38

List of Figures

1.1	Typical microgrid system [1].	2
2.1	Machine Learning Regression Flowchart.	11
3.1	Silk house Bragança [2]	15
3.2	Microgrid implemented on Silk house [3]	16
4.1	Study site emplacement.	20
4.2	Weather data.	20
4.3	Data density.	21
4.4	Box and whiskers plot.	21
4.5	Humidity data processing	22
4.6	Wind speed data processing	22
4.7	Data correlation without considering outliers removal	23
4.8	Data correlation considering wind speed data outliers removal	24
4.9	Data correlation considering relative humidity data outliers removal	24
4.10	Data correlation considering relative humidity and wind speed outliers removal	25
5.1	Adopted forecasting principle	26
5.2	Models accuracy comparison considering the root mean squared error	30
5.3	Residuals plot for exponential GPR	30
5.4	Regression plot of fine Gaussian SVM	31
5.5	Regression plot of medium tree	32
5.6	Regression plot of medium tree	33
5.7	Regression plot of trilayered neural network	34
5.8	Regression plots for exponential GPR.	35
5.9	Solar Radiation testing using GPR Model	35

List of Tables

- 4.1 Weather data information. 21
- 5.1 Training models evaluation without removing outliers 28
- 5.2 Training models evaluation data treatment 29

List of abbreviations

CO₂	Carbon Dioxide
DG	Distributed Generators
EMS	Energy Management System
EV	Electric vehicles
GPR	Gaussian Process Regression
Kw	kilo Watt
LV	Low voltage
MAE	Mean Absolute Error
MG	Microgrid
ML	Machine Learning
MSE	Mean Squared Error
MV	Medium voltage
MWh	Mega Watt Hour
MW	Mega Watt
NN	Neural Network
PV	Photovoltaic
RES	Renewable Energy Sources

RH	Relative Humidity
RMSE	Root Mean Square Error
SVM	Support vector machines
WT	Wind Turbine

List of symbols

R^2	R-Squared
-------	-----------

Chapter 1

General Introduction

1.1 Introduction

A microgrid is a small power distribution system that contains loads and distributed energy sources (DERs) (such as distributed generators, storage devices, or controllable loads and non-controllable loads) (Figure 1.1) and can operate in a controlled and coordinated manner when connected to the main grid or islands [4].

A microgrid is a self-sufficient energy system that serves a discrete geographic footprint, such as a university campus, hospital complex, commercial center, or community (15). This type of storage devices can be a farm of batteries often from batteries. Some now also have electric vehicles (EV) charging stations. Microgrids are interconnected with nearby buildings, providing customers with electricity, possibly heating and cooling, through Energy Management System (EMS) and control systems [5].

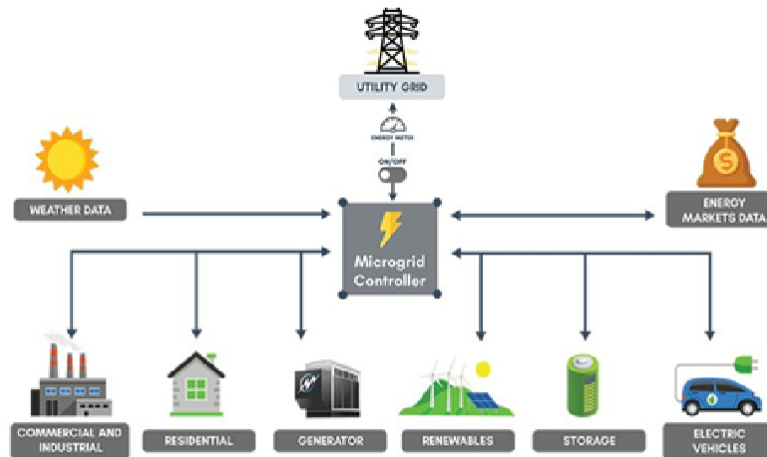


Figure 1.1: Typical microgrid system [1].

1.2 Motivations

Information about global solar radiance is the basis for many solar energy uses as well as economic and environmental considerations. However, because solar radiance changes and measurements are sometimes unavailable, it is often difficult or impossible to obtain accurate global solar radiance data. Machine learning models, on the other hand, are capable of solving highly nonlinear problems. They have many potential applications and are of great interest to researchers worldwide. Particularly, regression machine learning models are a best tool to perform solar radiance prediction in short and long term [6]. However, in term of microgrid, solar radiance forecasting can assist management of an isolated microgrid powered by photovoltaic systems (PV) and equipped with an energy storage system (ESS). Accurate forecasting is needed to reduce the uncertainty of photovoltaic array caused by solar radiance intermittency and control optimally the microgrid: schedule production, storage and load to maximize renewable resources production, ESS lifetime and minimize the customer energy costs. The energy management system (EMS) must take into account the uncertainty of the forecasting models to elaborate control strategies. The short-term solar forecasting is coupled

with an auto-regressive model.

1.3 Literature Review

Machine learning (ML) is a sub-field of computer science and it is classified as an artificial intelligence method [7]. It can be used in several domains and the advantage of this method is that a model can solve problems which are impossible to be represented by explicit algorithms [8]. Many methods have been developed to predict solar radiation, including theoretical parameter models, empirical models, artificial intelligence models, and satellite retrieval data [9]. In recent years, several works have been performed to forecast the solar resource employing Machine Learning techniques and environmental parameters. The development of machine learning has inspired many researchers to use machine learning algorithms to develop solar radiation prediction models. Several authors have used the Artificial Neural Network (ANN) models [10, 11], some researchers have carried out the comparative analysis of a variety of machine learning algorithm and all these works show that the ANN algorithm does not realize good prediction results [12]. Other approaches have considered generalized fuzzy models [13]. In [14] the authors used the Extreme Learning Machine (ELM). support vector machines (SVMs) have been also widely applied to forecast solar radiance in short as the work presented on the paper [15]. Tree algorithms, such as the random forest algorithm and the gradient boosting regression tree (GBRT) algorithm, have been used in [16] to construct solar radiation prediction models and shown an interesting results . Gaussian process regression (GPR) in [17] shown an excellent result to forecast solar radiance, GPR is simpler and generally more robust than other statistical regression tools. In addition to the good numerical performance and stability, GPR requires a relatively small training data set, it can adopt very flexible kernel functions, and it provides confidence intervals for the predictions.

1.4 Objectives

The work presented in this thesis has several objectives mainly :

1. Analysis and weather data processing using statistical Approaches.
2. Solar irradiation prediction using supervised machine learning models based on regression including: linear regression models, regression trees, Gaussian process regression models, support vector machines, kernel approximation, ensembles of regression trees, and neural network regression models.

3. Evaluation of Machine learning training and testing results using the Root Mean Squared Error (RMSE), Mean Square Error (MSE), R-squared, Mean Absolute Error (MAE).
4. Exporting the best obtained regression model and testing it with new data-set to assess the testing performances.

1.5 Thesis Structure

The work presented in this thesis is organized into six chapters :

Chapter 1 (General Introduction), addressed to show the general idea of the problem, the motivations that led to its treatment, the expected objectives of the thesis and the structure followed in its writing.

Chapter 2 (State of the Art) aims to describe the renewable energy technologies and their use inside microgrid systems equipped with energy management system, furthermore, it explain generally the regression machine models and methods of their evaluation, and precisely, present the machine learning regression principle adopted on this thesis.

Chapter 3 (Silk house Building) present the Silk house building and its microgrid infrastructure.

Chapter 4 (Case of study) define the case of study and present the weather data processing.

Chapter 5 (Results and Discussion) contains the results achieved by the implementation regression machine learning results based on regression learner toolbox of *MATLAB* Apps, the results are interpreted and justified in order to finally export the best obtained model based on training and testing errors evaluation.

The last chapter conclude the study and propose guidelines for future works.

Chapter 2

State of The Art

2.1 Introduction

The solar radiation is an important parameter for solar research, but it is not available at most sites due to the lack of solar radiation meters at weather stations. Therefore, it is important to predict insolation at a location based on multiple climate variables. These variables are sunshine duration, maximum ambient temperature, relative humidity, latitude, longitude, day of year, daily clear sky global radiation, total cloud cover, temperature, clarity index, altitude, month, mean temperature, mean Cloud cover, mean wind speed, atmospheric pressure, reference clarity index, mean diffuse radiation, mean radiation, monthly, extraterrestrial radiation, evaporation and ground temperature. Sunlight duration is easily obtained and measured in most locations, so it is often used to simulate solar radiation [7].

2.2 Renewable Energy

An energy framework is comprised of an energy supply area and energy end-use advances. The object of the energy framework is to convey to shoppers the advantages that energy offers. The term energy administrations are utilized to depict these advantages, which in families incorporate enlightenment, prepared food, agreeable indoor temperatures, refrigeration, and transportation. Energy administrations are additionally needed for all intents and purposes each business and modern movement. For example, warming and cooling are required for some modern cycles, thought process power is required for farming, and power is required for media communications and hardware [18]. The energy chain that conveys these administrations starts with the assortment or extraction of essential energy that, in one or a few stages, might be changed over into energy transporters, for example, power or diesel oil, that are appropriate for

end employments. Energy end-use hardware—ovens, lights, vehicles, apparatus—changes over conclusive energy into helpful energy, which gives the ideal advantages: the energy administrations. An illustration of an energy chain—starting with coal extraction from a mine (essential energy) and finishing with delivered steel as an energy administration. Energy administrations are the consequence of a mix of different advances, framework (capital), work (expertise), materials, and essential energy. Every one of these sources of info conveys a sticker price, and they are incompletely substitutable for each other. According to the purchaser's viewpoint, the significant issues are the monetary worth or utility got from the administrations. Buyers are frequently ignorant about the upstream exercises needed to create energy administrations [18].

2.3 Renewable Energy Technologies

Renewable energy sources (counting biomass, sunlight based, wind, geothermal, and hydropower) that utilization native assets can possibly furnish energy administrations with nothing or very nearly no outflows of both air poisons and ozone depleting substances. Presently, environmentally friendly power sources supply 14% of the all-out world energy interest. The stockpile is overwhelmed by conventional biomass utilized for cooking what's more warming, particularly in rustic areas of agricultural nations. Huge scope hydropower supplies 20% of worldwide power. Its extension for development is restricted in the industrialized world, where it has almost arrived at its financial limit. In the creating scene, significant expected still exists, however huge hydropower tasks might confront monetary, ecological, and social limitations. Through and through, new environmentally friendly power sources contributed 2% of the world's energy utilization in 1998, including 7 exajoules from current biomass and 2 exajoules for any remaining renewables (geothermal, wind, sun based, and marine energy, and limited scope hydropower). Sun oriented photovoltaics and framework associated breeze introduced limits are developing at a pace of 30% per year. All things being equal, it will probably be a very long time before these new renewables amount to a significant part of all out-energy utilization, since they as of now address such a little rate [18].

2.3.1 Photovoltaic Power

Photovoltaic sun-based energy is gotten from the transformation of daylight into power inside semiconductor materials, for example, silicon, which are covered with a slim metal layer. These photosensitive materials have the property of delivering their electrons affected by outer energy, the energy is conveyed by photons, which is a principal part of light that hit the electrons and delivery them, accordingly, an electric flow is made, then, at that point, this created

power is either accessible in direct structure or put away in batteries (decentralized electrical energy) or in power infused into the matrix, all relying upon the sort of plan and for which the establishment is planned. The photovoltaic sun-based generator is comprised of photovoltaic modules themselves made from photovoltaic cells associated with one another. The presentation of a photovoltaic framework relies upon the direction of the sunlight powered chargers and the insolation zones in which they are found [19].

2.3.2 Wind Energy

The primary conceivable source is wind energy. Perfect and sustainable, it is accessible in bountiful amounts all around the planet. As of late, wind power has turned into the principal wellspring of power from sustainable sources, barring hydro-power, with in excess of 593 TW h delivered worldwide in 2019. The change of this energy into power is finished by wind turbines, which can cover an enormous scope of force contingent upon the necessities and the kind of age picked. For instance, seaward wind turbines can be many meters high to produce a few MW, different kinds of lower stature that don't surpass several meters are adjusted for homegrown use, found near the puts in of supply in request to make hundreds to thousands of watts [20]. The wind energy is an endless asset, free and accessible wherever depicted by the development of barometrically air delivered by the non-uniform sun powered radiation on the outer layer of the earth, the sun warms the earth unevenly, which makes zones of various temperatures and environmental strain from one side of the planet to the other, from these tension contrasts, air developments are conceived, this last option addresses the breeze. Assuming a hotness assessment of the globe is made, an overflow of energy at the equator and a shortage at the shafts can be noticed, which prompts an exchange of hotness from the equator to the posts, so the mass of warmed air at the equator chills off in the shafts. As a result, this removal will bring about the breeze, it is depicted by the laws of liquid mechanics. The earth gets through the sun a force of 1.74×10^{17} W simply 1 to 2% of this energy is changed over into wind energy. Notwithstanding, wind variety in each site is a significant rule that should be painstakingly viewed as while carrying out a wind turbine framework [21].

2.3.3 Hydroelectric Power

Here is an overall view that hydroelectricity is the environmentally friendly power source standard excellence—no exhaustible, nonpolluting, and more monetarily appealing than different choices. Hydropower plants radiate substantially less ozone-depleting substance than do warm plants. Ozone-depleting substance emanations of hydropower are brought about by the rot

of vegetation in overwhelmed regions and by the broad utilization of concrete in the dam development. Sadly, there are neighborhood effects of the utilization of waterways, social just as biological, and they are acquiring significance as individuals become mindful of what those effects mean for expectations for everyday comforts. Then again, most sustainable wellsprings of energy (like hydroelectricity age) are capital concentrated however have lower functional expenses than warm and atomic choices. The high introductory expense is a not kidding boundary for its development in emerging nations, where a large portion of the undiscovered monetary potential is found [22].

2.4 Microgrid Management System

The utilization of electric energy assumes a vital part in the financial improvement of the population; hence, there is a developing energy interest in both metropolitan and country regions all over the planet, which should be fulfilled by expanding the age limit. The utilization of electric energy has introduced a huge increment, considering the size of the population, the sped-up modern development and different variables that straight forwardly affect this peculiarity. For the year 2018, electric energy utilization of north of 69 GWh was acquired, with a development of more than 3% regarding the earlier year. Not with standing, a few provincial regions don't have energy supply, so it is needed to stretch out the power network to these spots, or when this choice is impossible, the execution of nearby age frameworks ordinarily made from warm generators dependent on petroleum derivatives is done. Starting at 2014, roughly 61% of the energy produced on the planet came from consuming petroleum derivatives (flammable gas and oil). In any case, these energies are the primary driver of ecological contamination, which is producing environmental change around the world, as well as burning-through restricted normal assets that eventually will be inadequate to satisfy energy need. To such an extent, that constantly 1990 the yearly increment of CO_2 fixations was 0.44% while from the year 2000 this normal expanded to 0.55% each year, arriving at 400.8 parts per million out of 2016, the most elevated focus since the start of the modern time. In this way, it has turned into a need to utilize environmentally friendly power sources in power age as a choice to battle environmental change. The momentum needs to foster manageable systems to meet the energy prerequisites of non-interconnected regions has been quite difficult for the two specialists and legislative substances. In this manner, microgrids arise as conceivable energy arrangements, especially for regions not interconnected to the power matrix, which limit the ecological effect of age frameworks by including environmentally friendly power sources accessible nearby, for example, water, wind, sun-based radiation and biomass. A microgrid is

a coordinated energy framework comprising of interconnected loads and disseminated energy assets, connectable and island able, portrayed by being profoundly discernible and controllable, being able to give capacity to little electrical burdens, whose size can be variable (by and large it is considered a microgrid if the size is under 500 kW and micro-system assuming the size is under 50MW). In like manner, Smart Grids are incorporated to microgrids through bidirectional correspondence processes between specialist co-ops and clients associated with the framework. These are completely coordinated and mechanized frameworks with the goal of supplanting customary constructions and adding to the development of the electric lattice and expanding its effectiveness, just as a reaction to the need to modernize the electric matrix, articulating control and checking processes with green innovations. Microgrids arise as a framework that helps the activity and inclusion of force frameworks for creating economies. They are outlined in the disseminated age climate, where the energy sources utilized can be both inexhaustible (wind generators, photovoltaic and photo-thermal innovation, biomass, and so on) and non-sustainable (diesel or petroleum gas generators, power devices, and so forth), every one of them cooperating with capacity hardware for example, capacitor banks or batteries. In this manner, the plan of secluded microgrids should be done considering the administration and reconciliation of the distinctive age and capacity frameworks, to fulfill the need and nature of administration prerequisites. In the writing there are a few plan approaches that try to meet this evenhanded; some of them depend on the utilization of improvement techniques, others use recreation devices or probabilistic examination of the factors engaged with the plan. Among the fundamental open doors that microgrids present in the energy market, it is critical to feature that they suggest an increment in energy productivity, lessen CO_2 emanations, increment the infiltration of renewable and diminish the energy cost. Also, since age is done locally, energy misfortunes in huge dissemination lines would be decreased. As a supplement to the natural benefits. Microgrids are not another idea. If the meaning of microgrid is applied in a real sense, Thomas Edison's first electrical power age office on Pearl Street in Manhattan was the first microgrid in history. Be that as it may, the real microgrid idea was initially evolved to disengage and keep a low voltage dispersion network in case of shortcomings upstream of the LV/MV transformer. Similarly, in case of a summed-up power disappointment, microgrids could aid the rebuilding system [23].

2.5 Machine Learning

Machine learning (ML) is the study of computer algorithms that can improve automatically through experience and using data. It is seen as a part of artificial intelligence. Machine

learning algorithms build a model based on sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, speech recognition, and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks [24]. A subset of machine learning is closely related to computational statistics, which focuses on making predictions using computers; but not all machine learning is statistical learning. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning. Some implementations of machine learning use data and neural networks in a way that mimics the working of a biological brain. In its application across business problems, machine learning is also referred to as predictive analytic [24]. In the scope of performing energy management of the microgrid system to further perform system improvement regarding the site equipment, it is required to perform an energy forecasting of the weather parameters including, Solar radiation. By training machine learning models using the existed database that is described on the next chapter. Weather forecasting parameters is a wide research topic that has been explored by researchers. In the work presented on this thesis, the weather data parameters have been performed based on regression machine learning methods including Linear Regression, Stepwise Linear, Decision Trees, Support Vector Machine (SVM), Gaussian Process Regression (GPR) and Neural Network (NN). According to work need, solar radiation were predicted using the previously recorded values of each parameter, temperatures and relative humidity and wind speed as predictors, then the best trained model is explored to evaluate the microgrid performances and provide improvement of the system if required.

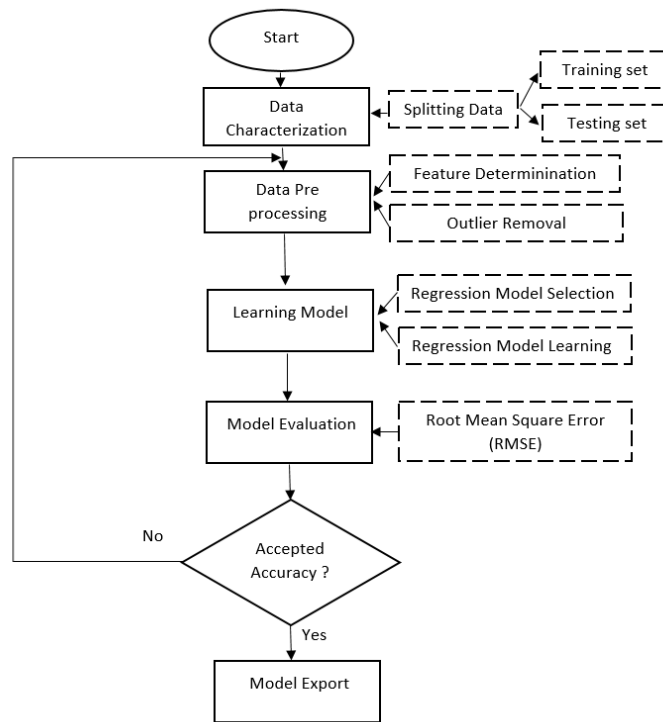


Figure 2.1: Machine Learning Regression Flowchart.

2.6 Errors

In this work the machine learning errors were evaluated by: mean square error , root mean square error, R-squared and mean absolute error.

2.6.1 Mean Square Error

The mean squared error (MSE) lets you know how close a relapse line is to a bunch of focuses. It does this by taking the good ways from the focuses to the relapse line (these distances are the "errors") and figuring out them. The squaring is important to eliminate any regrettable signs. It likewise gives more weight to bigger contrasts. It's known as the mean squared error as you're tracking down the normal of a bunch of blunders. The lower the MSE, the better the forecast [25].

Formula of MSE:

$$MSE = \frac{\sum_{i=1}^n (Y_i^r - Y_i^p)^2}{n} \quad (2.1)$$

where:

- Y_i^r is the actual value.
- Y_i^p is the predicted value.
- n is the number of observations.
- MSE is the mean square error.

2.6.2 Root Mean Square Error

Root Mean Square Error (RMSE) is the standard deviation of the residuals (forecast blunders). Residuals are a proportion of how a long way from the relapse line information focuses are; RMSE is a proportion of how fanned out these residuals are. As such, it lets you know how focused the information is around the line of best fit. Root mean square error is regularly utilized in climatology, gauging, and relapse investigation to check trial results [26].

The formula is:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{i=1}^n (Y_i^r - Y_i^p)^2}{n}} \quad (2.2)$$

2.6.3 R-squared

R-squared R^2 is a factual measure that addresses the extent of the fluctuation for a reliant variable that is clarified by a free factor or factors in a relapse model. Though connection clarifies the strength of the connection between an autonomous and ward variable, R-squared discloses how much the change of one variable clarifies the difference of the subsequent variable. In this way, on the off chance that the R^2 of a model is 0.50, around half of the noticed variety can be clarified by the model's bits of feedbacks [26].

Formula of R-Squared is:

$$R - Squared = \sqrt{\frac{\sum_{i=1}^n (Y_i^r - Y_i^p)^2}{\sum_{i=1}^n (Y_i^r - \hat{Y}_i)^2}} \quad (2.3)$$

- \hat{Y}_i Mean of all actual values.

2.6.4 Mean Absolute Error

Mean Absolute Error is a model assessment metric utilized with relapse models. The mean outright error of a model as for a test set is the mean of the outright upsides of the singular forecast blunders on over all cases in the test set. Every expectation error is the contrast between the genuine worth and the anticipated incentive for the case [27]. Formula of MAE:

$$MAE = \frac{\sum_{i=1}^n (Y_i^r - Y_i^p)^2}{n} \quad (2.4)$$

Chapter 3

Silk house Building

3.1 Introduction

Silk house is a museum housed in an old mill on the left bank of the Fervença River in the center of Bragança (Figure 3.1). This is an interactive space, part of the Ciência Viva Center in Bragança, dedicated to the history of silk. Silk house is housed in an old building where silk was dyed in the 18th century. During the 19th and 20th centuries, the property was used as a mill and still maintains the original character of its building. Later, it became private property, using the mill's potential to produce regional bread and biscuits. In 1990, the municipality of Bragança acquired the building and restored it during the intervention of the Polis plan in 2006. Today, Silk house is a museum dedicated to the dissemination and dissemination of science, where courses, lectures and exhibitions are held. The museum also has a permanent exhibition on the history of silk that outlines the various stages of silk production, from caterpillar development and cocoon formation to the spinning and weaving process. Silk house receives an average of more than 11,500 visitors per year. The building is in the city center of Bragança, at an altitude of 664 m. The building is a terraced building with a total of 3 floors and is full of sunlight. It features a stone structure, wooden roof and tiles, plastered walls, wooden openings with clapboards and original traditional stone masonry [2].



Figure 3.1: Silk house Bragança [2]

3.2 Electrical Characterization of the Building

The power installation is powered by a low-voltage three-phase distribution network with a contract power of 13.6 kW. This is a weekly building that operates from Tuesday to Sunday, centered during the day, between 10:00 and 18:00. In addition to normal operations, various events are held regularly, especially on weekends, extending events until around midnight [2].

3.3 Characterization of the Energy Potential of the Building

From the investigation of the building, 2 structures stood out for the study of its energy potential: the rooftop with superb sun openness and the construction of the old factory and an embankment right close to Silk house. Considering the structure's rooftop region, there are two favored regions for the establishment of photovoltaic modules: one with $40 m^2$ and the other with $20 m^2$. Both are pointing toward the south and have a rooftop pitch of around 16° . The yearly creation capability of photovoltaic energy is around 1395 kWh/kWp each year. As indicated by genuine information from photovoltaic units introduced starting around 2010, in an area with comparable climatic qualities, 2 km away, on the IPB Campus, the energy creation is 1466.9 kWh each year, per introduced kWp., for this situation the slant was upgraded for the area [2].

3.4 Microgrid in Silk house

This segment presents the brilliant microgrid created by the above necessities. The microgrid coordinates appropriated age (PV and pico-hydro), energy stockpiling considering the battery bank, the executives, and observing framework. It is three-stage (400 V, 50 Hz) microgrid to supply the House of Silk loads. The shrewd microgrid depends on the SMA Flexible Storage System with battery-reinforcement and expanded of self-utilization. It is created by three Sunny Island 4.4M, three PV inverters Sunny Boy 1.5, one Sunny Home Manager 2.0, two SMA Energy Meter, three strings of six Panasonic's PV modules VBHN325SJ47, one 300W level water wheel to recuperate the authentic legacy of the previous plant, one low-head turbine (LH400 from Power Spout), one PV microinverter, one PV inverter, a battery bank comprising of 24 VRLA 2 V battery's squares Sonnenschein's batteries A602/625 Solar, battery intertwine, electrical board, programmed move switch, and securities. Figure 3 presents a circuitry outline of the brilliant microgrid carried out in the House of Silk. The activity of the microgrid can be portrayed as follows. The three-stage microgrid is the outer electric lattice at whatever point it is accessible. On the off chance that the instance of matrix disappointment, the microgrid is set up by the three Sunny Island (SI) 4.4M associated in an expert slave setup. A programmed move switch is dependable by the outside framework association or detachment. The energy created by the PV strings is infused into the microgrid, distributing the power by the three stages utilizing three Sunny Boy 1.5 [3].

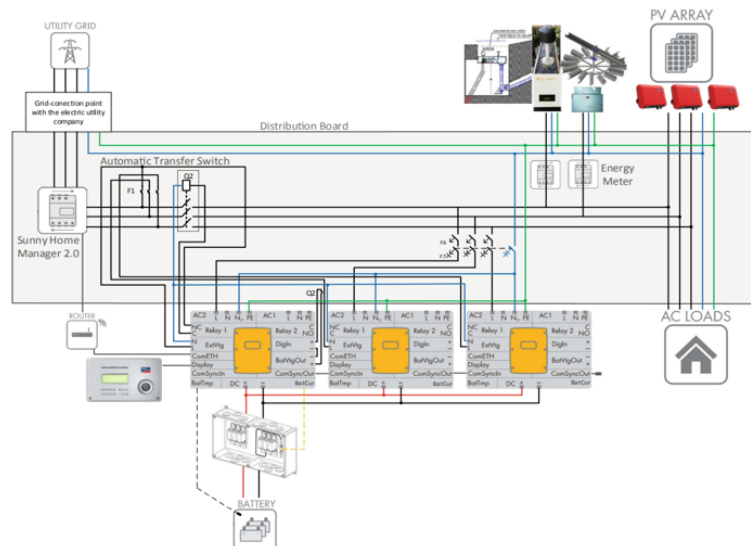


Figure 3.2: Microgrid implemented on Silk house [3]

At the point when the age is higher than the utilization, the overabundance of energy is

put away in the battery bank by the three bi-directional battery inverters, SI. These inverters utilize the batteries to control the power stream and, in this way, to further develop self-utilization. The dynamic power produced by the inverters associated with the PV strings is restricted by recurrence control, as better clarified later. This happens at whatever point the age is higher than utilization, the batteries are charged, and the energy took care of into the outside framework should be zero. At the point when the utilization is higher than the age, the energy comes from the batteries and, assuming it sufficiently isn't, the excess energy is bought from the outer network to take care of the heaps. If there should be an occurrence of network disappointment, the microgrid is disconnected from the matrix utilizing the programmed move switch and starts working in remain solitary mode. In this activity mode, a few hours of independence are accessible, yet it firmly relies upon the utilization, the age, and the battery bank condition of charge. For this situation, a few burdens can be naturally separated [3].

3.4.1 Photovoltaic system

The photovoltaic system comprises of three strings with six PV modules each and was introduced on the rooftop with a slant of 12° and two shadowless regions. The first is situated South and the second is arranged Southwest. Exploiting this, the PV strings were separated by these two rooftop surfaces to work on the dispersion of PV age for the duration of the day. One PV string is situated South and the other two are arranged Southwest. In the first part of the day, the Southern arranged string begins the age before than the Southwestern situated ones, permitting a higher creation. In the early evening, the Southwestern arranged strings will expand their age longer. These directions make conceivable to work on the appropriation of photovoltaic age for the duration of the day, changing it to the opening times of the mesum [3].

3.4.2 Pico-Hydro power

The Silk house electric age will be supplemented, principally in the colder time of year terms, by two pico-hydro frameworks that are being introduced. These frameworks are limited scale hydropower age units up to 5 kW that changes over the force of streaming water of a waterway, stream or stream in power. These low power frameworks, however equipped for delivering energy 24 h daily, in contrast to what occurred until a couple of years prior, it is feasible to utilize super durable magnet simultaneous generators, with variable speed, contingent upon the stream rate and head. This happens on the grounds that inventive outcomes show that the energy created can be tackled utilizing PV inverters, deftly, and productively. Moreover, those generators and PV string inverters and miniature inverters are, these days, off-the-rack and broadly accessible

parts utilized in wind and photovoltaic applications. The establishment of the hydroelectric plant results from the nearness of Silk house of the waterway Fervença, the presence of a little dam (under 10 m) and the pressure driven foundation of the old plant, channels, and displays, which has been monitored. From one viewpoint, the embraced arrangement saves the structure's engineering and, on the other, it recuperates the structure's chronicled legacy by introducing a flat-water wheel to deliver energy instead of the previous factory. There are two exhibitions accessible. The first, at the upper level, was utilized to introduce the microgrid hardware. The other, where the factory was found, a 300W even water haggle 1.2 kW low-head turbine (LH400 from Power Spout) are being introduced. Two lines were introduced to catch water from the little dam and take it inside the display where the water haggle will be. The water is gotten back to the stream 18 m downstream. A PV microinverter from GWL Power and a PV string inverter from Omnik, model Omniksol - 1.5k - TL2 , are utilized to interface the 300W generator of the water haggle 1.2 kW generator of the low-head turbine to the microgrid, separately. After broad tests in the research center, the outcomes show that the inverters introducing the best presentation are not from SMA. Along these lines, SMA Energy Meters should be utilized to account the energy delivered by these inexhaustible assets in microgrid energy the board [3].

Chapter 4

Case of Study

4.1 Introduction

Precise forecasting of renewable energies such as solar is becoming one of the very important concerns in developing modern electrical grids. Hence, establishing appropriate tools of weather forecasting with satisfactory accuracy becomes an essential preoccupation in today's changing power world. The high variability of renewable sources makes forecasting crucial for maintaining the reliability of the grid. Improving the confidence level and the precision of forecasting is important for reducing the usual reserve capacity and will give rise to considerable savings in operating costs and capital. In this study we want to perform forecasting solar radiation using as predictor parameters: air temperature, relative humidity, and wind speed, as there's no data of silk house the chosen data are getting from the online platform of Malviya National Institute of Technology from Jaipur region in India (latitude: 26.8622°N, longitude: 75.8156°E) as shown on (Figure 4.1) including air temperature, wind speed, solar radiation and relative humidity as shown on the following figures. The exported data set is composed of 1922 features used for training the model. Once the required data from Silk house are collected, the best machine learning regression model will be applied for them. The goal of making this energy forecasting for the sensitive parameters is used for multipurpose in this thesis, first, performing an extended data base of solar radiation, in order to use it for performing an system sizing of the installation based on optimization algorithms, carrying out also an energy management for the designed microgrid system to further perform an energy analysis of the microgrid system, so that, according to the energy behavior of the microgrid system it can be decided concerning the architecture of the system or it should be reduced or extend regarding renewable energies equipment's.



Figure 4.1: Study site emplacement.

4.2 Data Characterization

In this study, real meteorological measurements were collected using a data monitoring system composed of a meteorological sensor for recording, the temperature, wind speed, solar radiation. The monitoring system can record the sensors' measurements with an interval of 15 minutes. The data are measured by the meteorological station of Malviya National Institute of Technology from Jaipur region in India (Latitude: 26.8622°N , Longitude: 75.8156°E) for the month of March 2017. The input samples represent a set of measurements quantified at 1922 values recorded each five minutes for both wind speed and temperature data. Following figures show the wind speed, temperature, wind speed and solar radiation data respectively measured each 15 minutes.

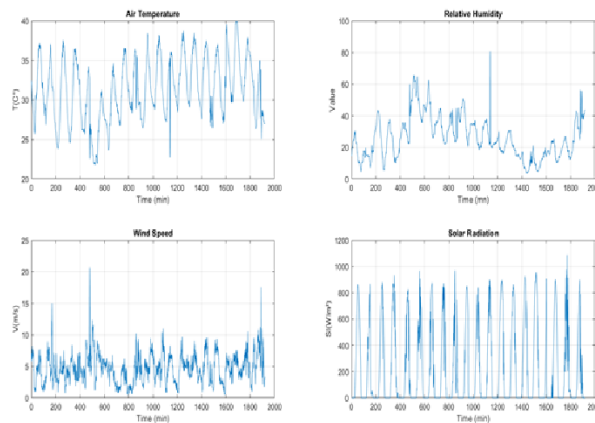


Figure 4.2: Weather data.

Table 4.1: Weather data information.

Parameters	Wind Speed (m/s)	Relative Humidity (%)	Temperature (C°)	Solar Radiation (W/m ²)
Min	0.47	3.41	21.75	0
Max	20.62	80.72	39.97	1089
Average	4.71	26.14	31.39	233.79

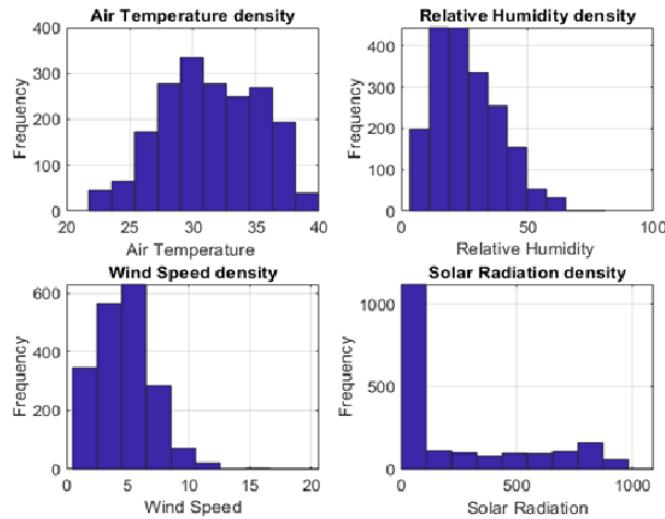
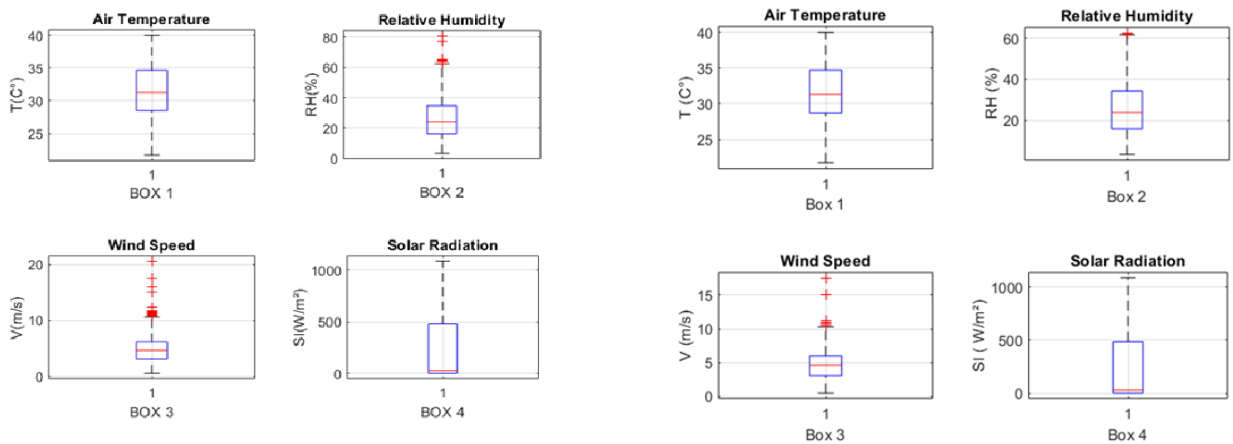


Figure 4.3: Data density.



(a) Before data treatment.

(b) After Data treatment

Figure 4.4: Box and whiskers plot.

Fig 4.5 and Fig 4.6 represent the humidity and wind speed data cleaning, before and after removing the outliers using the quartiles method.

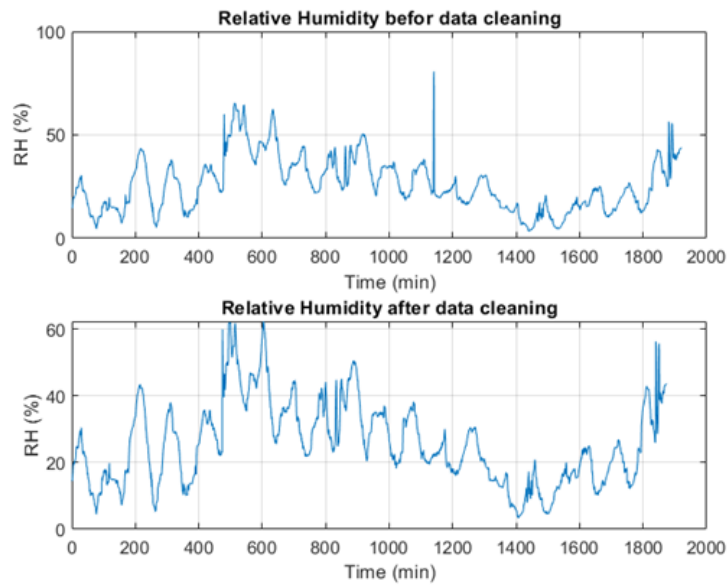


Figure 4.5: Humidity data processing

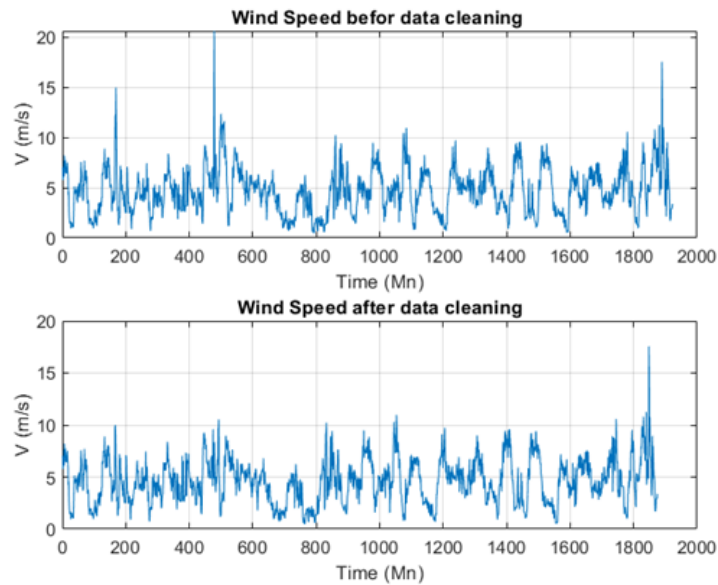


Figure 4.6: Wind speed data processing

4.3 Data Correlation

The general range of correlation is:

- Between 0 and 0.1 means that there no correlation.
- Between 0.1 and 0.3 means that there is a little correlation.

- Between 0.3 and 0.5 means that there is a medium correlation.
- Between 0.5 and 0.7 means that there is high correlation.
- Between 0.7 and 1 means that there is a very high correlation.

The Pearson correlation measures the strength of the linear relationship between two variables. It has a value between -1 to 1, with a value of -1 meaning a total negative linear correlation, 0 being no correlation, and + 1 meaning a total positive correlation.

In this section we have considered the correlation in the following way, in effect:

1. Data correlation without considering the data treatment (without data cleaning).
2. Data correlation considering the data treatment of the relative humidity.
3. Data correlation considering the data treatment of wind speed.
4. Data correlation considering the data treatment for wind speed and relative humidity.

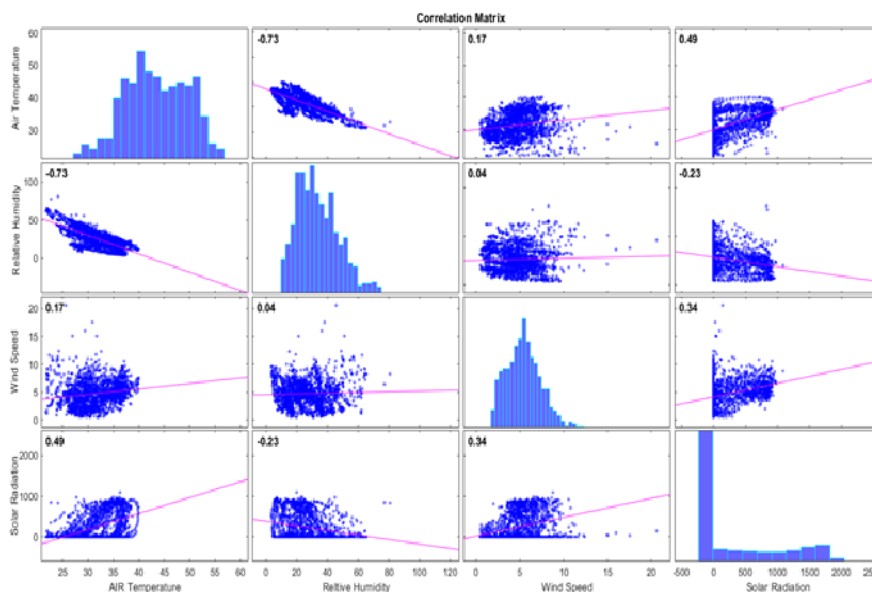


Figure 4.7: Data correlation without considering outliers removal

A high negative correlation is remarked between the air temperature and relative humidity(Figure 4.7), physically, the relation between humidity and temperature formula simply says they are inversely proportional. If temperature increases it will lead to a decrease in relative humidity(Figure 4.8), thus the air will become drier whereas when temperature decreases, the air will become wet means the relative humidity will increase. Due to the fact that air

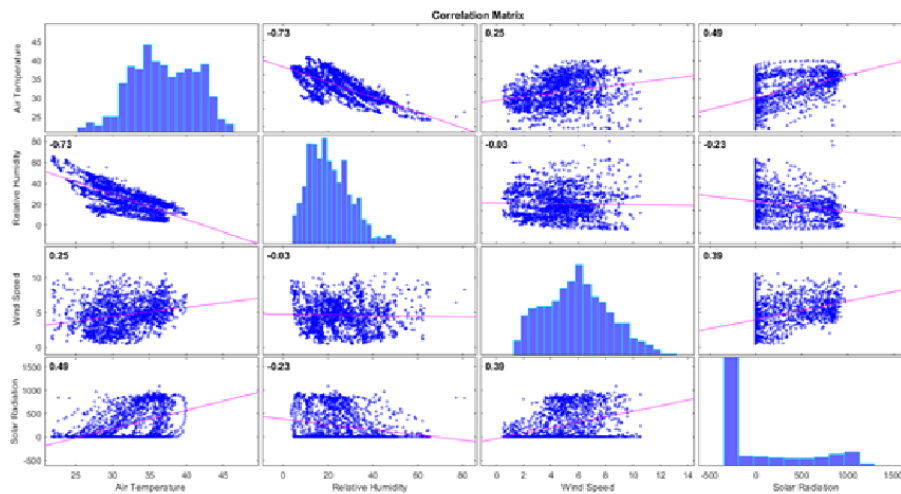


Figure 4.8: Data correlation considering wind speed data outliers removal

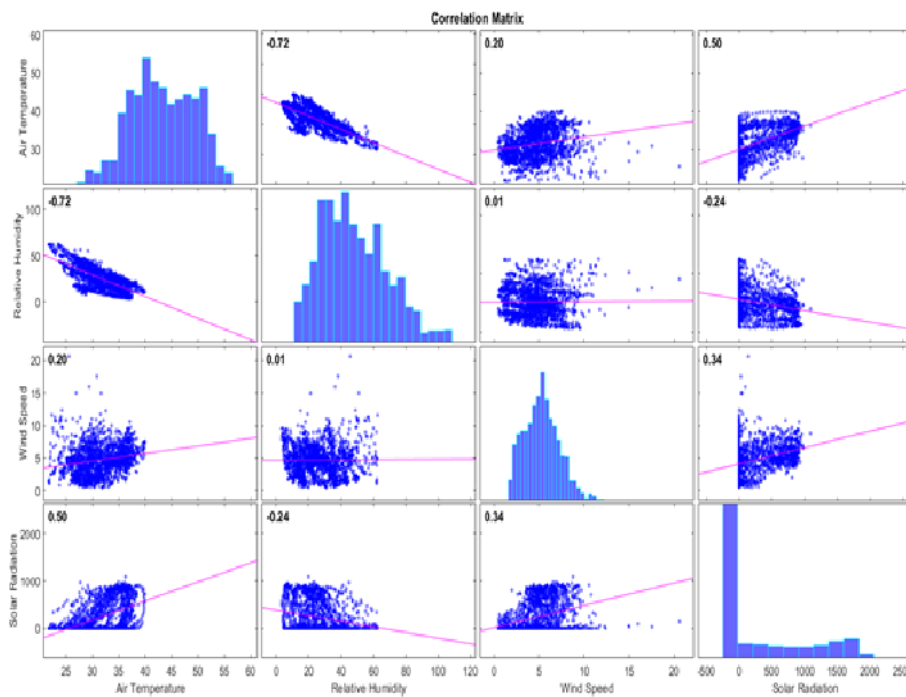


Figure 4.9: Data correlation considering relative humidity data outliers removal

temperature is directly related to the global solar radiation. So, an increase in solar radiation increases the air temperature and this justify the positive correlation gated between the two items(Figure 4.9). Because of the solar energy (incoming short-wave radiation), earth surface heats up. Air near the ground tends to gain heat from the earth surface. When air is heated, it expands and starts to rise. Solar energy causes wind due to its effect on air pressure for that

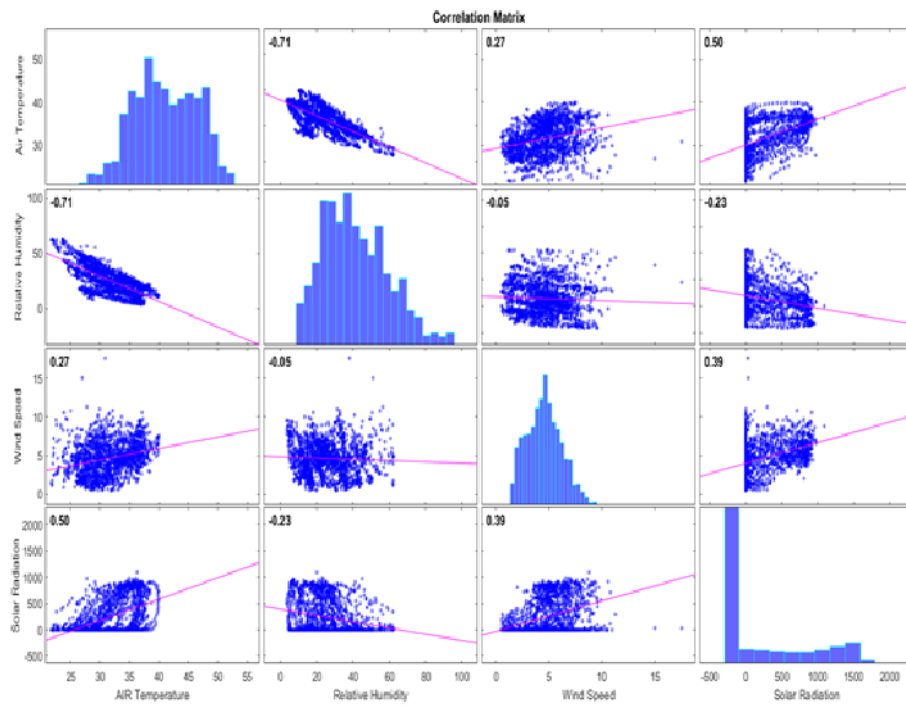


Figure 4.10: Data correlation considering relative humidity and wind speed outliers removal

there is a little correlation between them on the data analyzed. Non correlation is detected between relative humidity and wind speed (Figure 4.10) .

Chapter 5

Results and Discussions

5.1 Introduction

This chapter contains the simulation results using the regression learner toolbox from MATLAB software to forecast the solar irradiation. Two situations were considered, the first one the models are trained without considering the data processing in term of data cleaning by removing outliers, while, in the second situation, the data are treated before being used to train the models. to forecast the solar radiation, the models use as predictors: the time, the relative humidity, the temperature and wind speed. The forecasting principle is explained on the following figure:

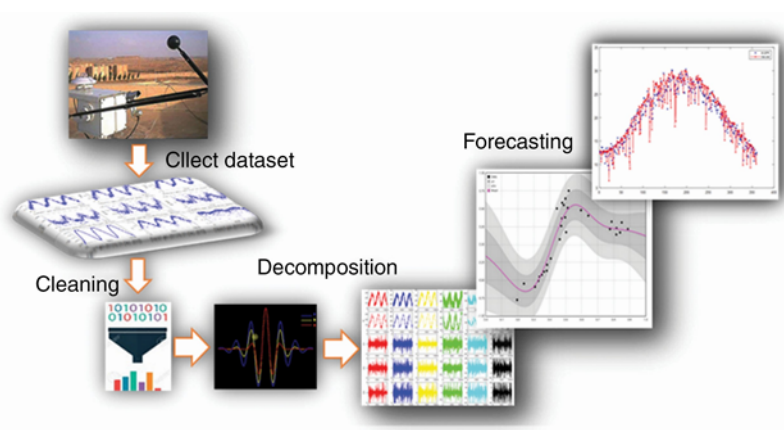


Figure 5.1: Adopted forecasting principle

5.2 Models Training and Validation

To discuss the training results, two situations were considered as follow.

5.2.1 Without Data Treatment

The first situation regards the data not treated in term of outlier removing, as presented previously, the relative humidity and wind speed are the parameters that contain perturbation however, to see their impact on model training, the data are used without pre-processing.

The following table represent the errors derived from the training of the regression models including linear regression models, regression trees, Gaussian process regression models, support vector machines, kernel approximation, ensembles of regression trees, and neural network regression models, using as predictors; wind speed, time, temperature and relative humidity. the errors considered were, the Mean Squared Error (MSE), the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), there definitions is presented on the previous chapter. Noting that the data set used to train the model is composed of 1189 value of sampling time of 15 between each values, this later were choose regarding their high correlation with the solar radiation, so, a changement if their behaviour can affect. In term of data training, the models were trained without cross validation for avoiding the under-fitting due to the short historical data base employed, The short set of data can justify the relatively elevated error of some models that need a big historical data set to make an accurate prevision. Inversely, the Gaussian Process Regression (GPR) had shown an excellent behaviour with a very small value of training error, offering for the regression model to be the best among the 24 trained models, the reason, is related to the architecture of the model in adoption of wave format data, as well his capacity to provide an accurate forecasting without giving refer to big historical training data set.

Table 5.1: Training models evaluation without removing outliers

Models	RMSE	R-Squared	MSE	MAE
Linear Regression (Linear)	244.97	0.37	60008	199.09
Linear Regression (Interactions Linear)	220.43	0.49	48591	174.24
Linear Regression (Robust Linear)	245.08	0.37	60063	198.59
Stepwise Linear (Stepwise Linear)	220.85	0.49	48773	173.41
Tree (Fine Tree)	61.111	0.96	3734.6	24.435
Tree (Medium Tree)	80.953	0.93	6553.3	36.265
Tree (Coarse tree)	102.78	0.89	10563	48.771
SVM (Linear SVM)	250.41	0.35	62707	197.28
SVM (Quadratic SVM)	168.84	0.7	28508	133.35
SVM (Cubic SVM)	140.63	0.79	19776	98.888
SVM (Fine Gaussian SVM)	59.142	0.96	3497.7	39.27
SVM (Medium Gaussian SVM)	109.96	0.87	12092	72.552
SVM (Coarse Gaussian SVM)	202.52	0.57	41016	153.31
Ensemble (Boosted trees)	87.927	0.92	7731.2	46.867
Ensemble (Bagged Trees)	75.595	0.94	5714.7	40.172
Gaussian Process Regression (Squared Exponential GPR)	74.415	0.94	5537.6	43.575
Gaussian Process Regression (Matern 5/2 GPR)	64.147	0.97	4114.8	36.323
Gaussian Process Regression (Exponential GPR)	0.1	1	0.010069	0.054
Gaussian Process Regression (Rational Quadratic GPR)	66.822	0.95	4465.2	37.862
Neural Network (Narrow Neural Network)	234.6	0.43	55039	178.96
Neural Network (Medium Neural Network)	80.669	0.93	6507.5	41.029
Neural Network (Wide Neural Network)	66.157	0.95	4376.8	39.177
Neural Network (Bilayered Neural Network)	87.755	0.92	7700.9	41.971
Neural Network (Trilayered Neural Network)	70.621	0.95	4987.4	35.127

5.2.2 With Data Treatment

This section concern the training of the models concerning the data treatment, by removing the data outliers from the model. In fact, the predictors are the same including, Relative Humidity, Temperature, Wind speed and Time. It can be remarked that the values of errors have been relatively reduced, this is due to removing of outliers that represented a failure values for the training models, the training is also performed without cross validation, The GPR model remain the best model in term of accuracy. Moreover, this model is going to be exported for being tested within testing data set composed of 100 samples with the same interval time sampling

as the one used for training.

Table 5.2: Training models evaluation data treatment

Models	RMSE	R-Squared	MSE	MAE
Linear Regression (Linear)	244.59	0.38	59822	199.15
Linear Regression (Interactions Linear)	221.05	0.49	48864	175.61
Linear Regression (Robust Linear)	244.66	0.38	59859	198.73
Stepwise Linear (Stepwise Linear)	221.46	0.49	49046	174.64
Tree (Fine Tree)	61.838	0.96	3823.9	25.086
Tree (Medium Tree)	79.669	0.93	6347.2	35.581
Tree (Coarse tree)	91.654	0.91	8400.4	42.917
SVM (Linear SVM)	249.02	0.36	62009	197.63
SVM (Quadratic SVM)	166.91	0.71	27857	132.05
SVM (Cubic SVM)	135.76	0.81	18431	97.168
SVM (Fine Gaussian SVM)	57.376	0.97	3291.9	39.206
SVM (Medium Gaussian SVM)	108.96	0.88	11872	72.008
SVM (Coarse Gaussian SVM)	201.2	0.58	40482	152.69
Ensemble (Boosted trees)	86.091	0.92	7411.6	47.065
Ensemble (Bagged Trees)	77.477	0.94	6002.7	42.258
Gaussian Process Regression (Squared Exponential GPR)	73.871	0.94	5456.9	43.447
Gaussian Process Regression (Matern 5/2 GPR)	63.679	0.96	4055.1	36.299
Gaussian Process Regression (Exponential GPR)	0.099	1	0.009	0.054
Gaussian Process Regression (Rational Quadratic GPR)	66.925	0.95	4478.9	38.242
Neural Network (Narrow Neural Network)	84.726	0.93	7178.5	42.309
Neural Network (Medium Neural Network)	78.398	0.94	6146.2	41.763
Neural Network (Wide Neural Network)	61.454	0.96	3776.6	38.42
Neural Network (Bilayered Neural Network)	78.985	0.94	6238.7	38.063
Neural Network (Trilayered Neural Network)	77.745	0.94	6044.3	36.568

Figure 5.2 represent the errors comparison in term of root mean squared error(RMSE) between the two models while considering and not considering the data cleaning, the difference is well remarked on the Narrow Neural Network.

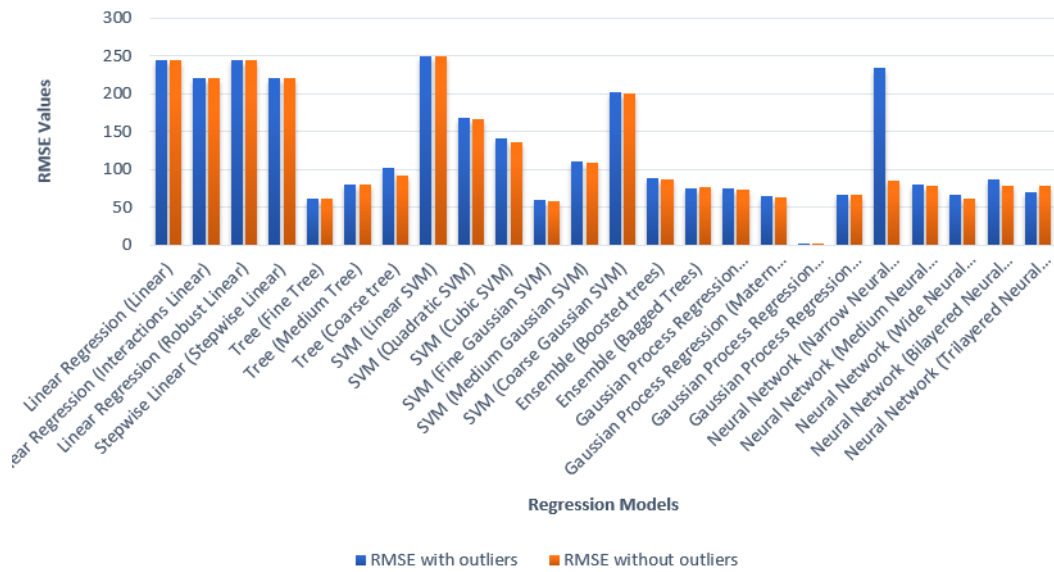


Figure 5.2: Models accuracy comparison considering the root mean squared error

Figure 5.3 is showing the residual plot for assessing the model training, according to the prediction results behaviour it can be shown that the residuals scattered roughly symmetrically around zero, that justifies the GPR model performances.

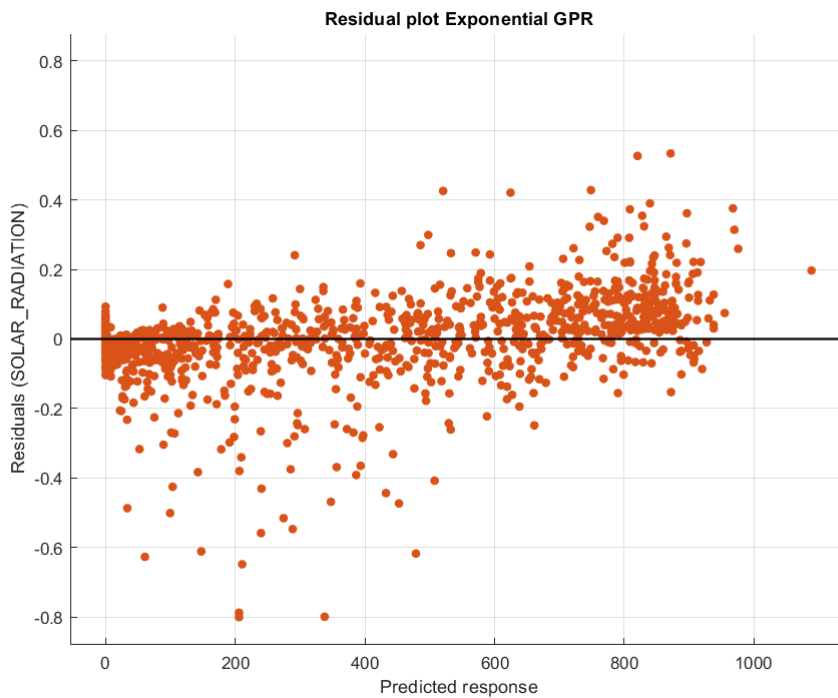


Figure 5.3: Residuals plot for exponential GPR

Figure 5.5 present the regression plot for model training of fine Gaussian SVM.

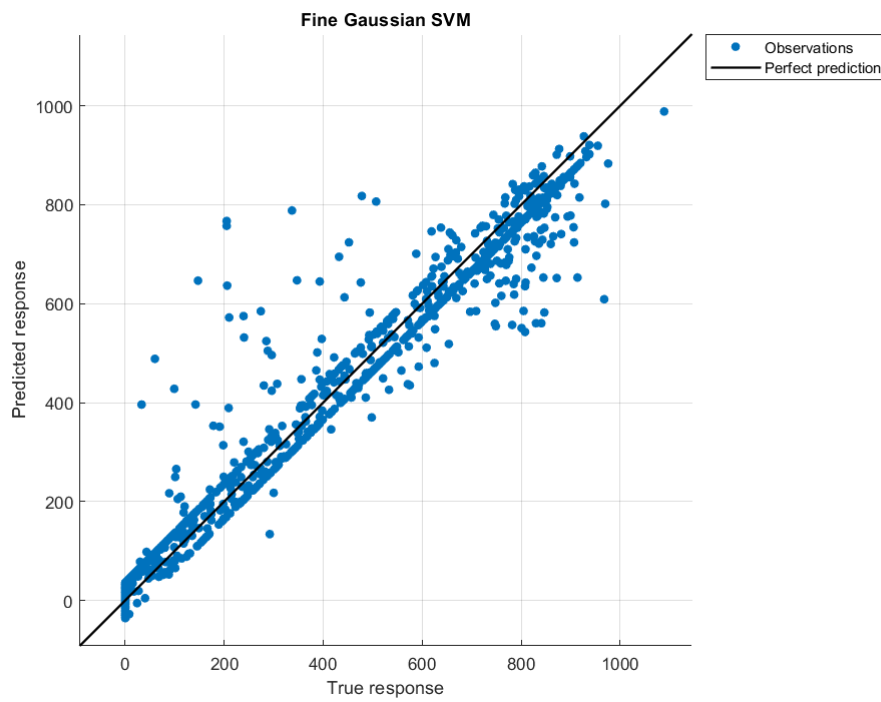


Figure 5.4: Regression plot of fine Gaussian SVM

Figure 5.5 present the regression plot for model training of medium tree.

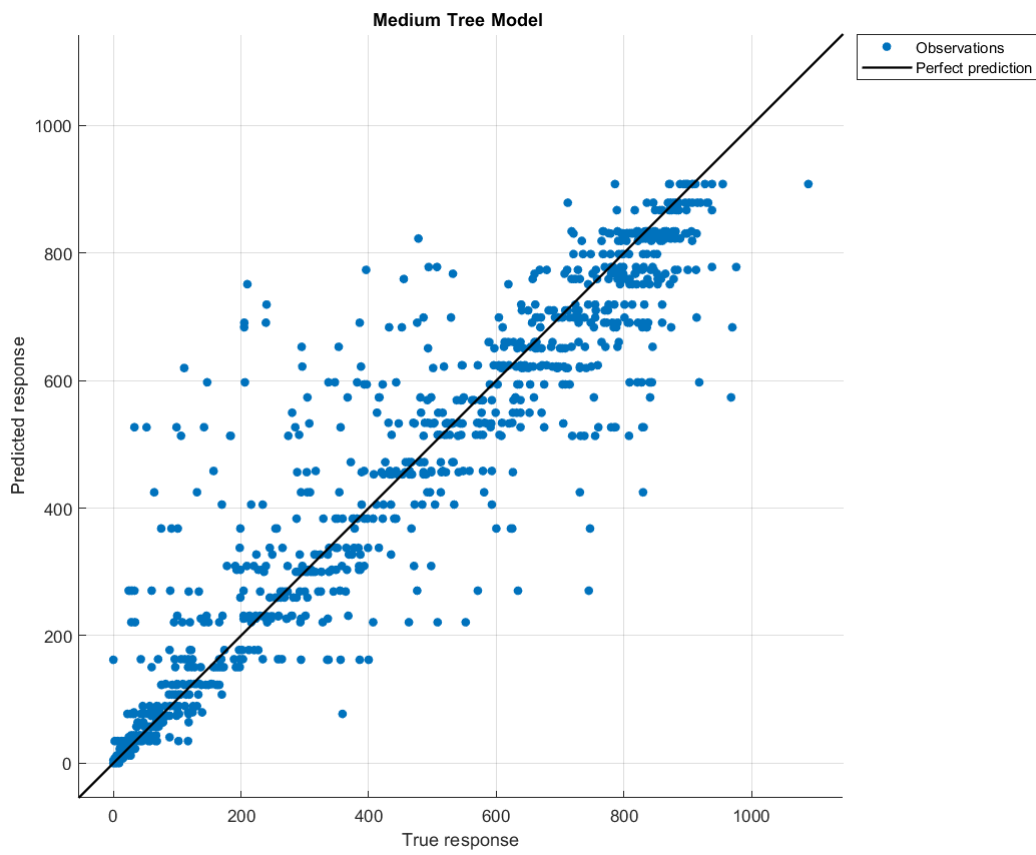


Figure 5.5: Regression plot of medium tree

Figure 5.6 present the regression plot for model training of medium tree.

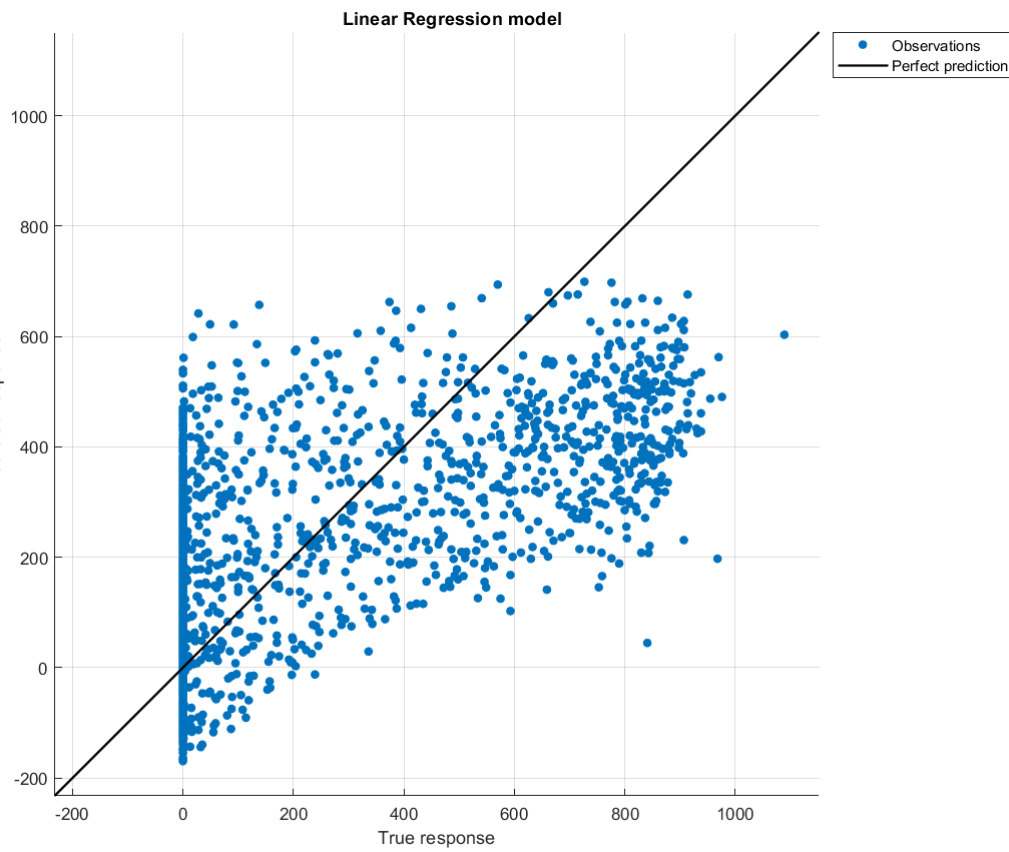


Figure 5.6: Regression plot of medium tree

Figure 5.7 present the regression plot for model training of trilayered neural network.

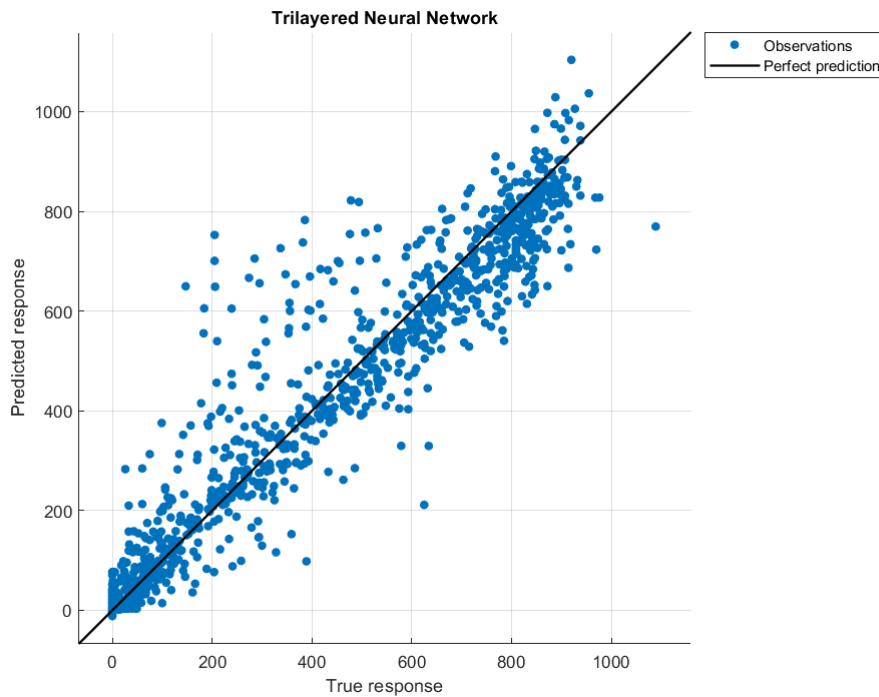
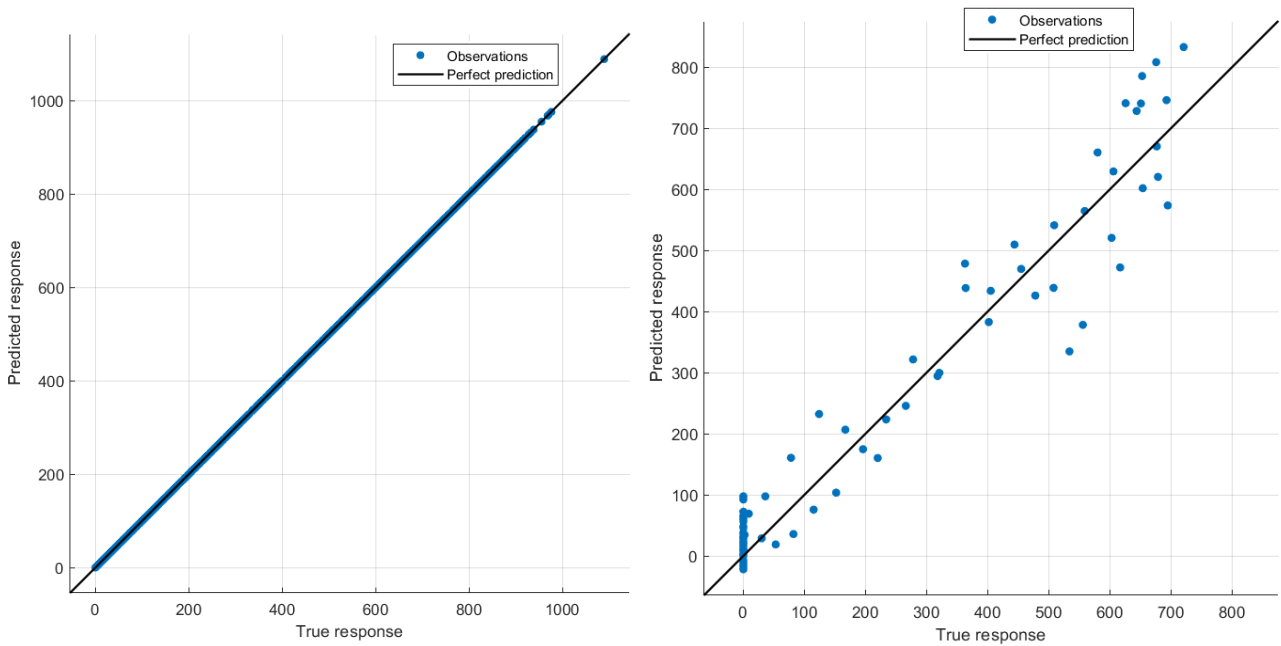


Figure 5.7: Regression plot of trilayered neural network

5.3 GPR model Exporting and Testing

Figure 5.8 is showing the predicted vs actual plot for the Gaussian process regression (Exponential GPR) for validation (Fig 5.8a) and testing data (Fig 5.8b). In order to test the model a data set of 98 samples have been used, the testing error evaluated by the RMSE was 23.02 . the diagonal represent the perfect prediction, in fact, the points are points scattered roughly symmetrically around the diagonal line, the vertical distance from the line to any point is small, this justify the small RMSE of 0.09 obtained by the exponential GPR model.

Figure 5.9 shows the solar radiance prediction results obtained after exporting the GPR model and tested with a set of 100 weather data set.



(a) Regression plot based on validation data

(b) Regression plot based on testing data

Figure 5.8: Regression plots for exponential GPR.

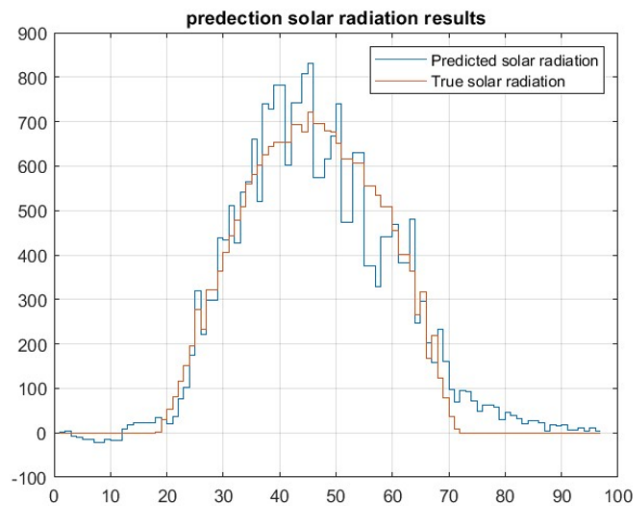


Figure 5.9: Solar Radiation testing using GPR Model

Chapter 6

General Conclusion and Perspectives

General Conclusion

The prediction of solar radiation is very important in the scope of sizing and optimal energy management of microgrid system, allowing to have a pre-decision regarding the system design avoiding an over or under infrastructures and avoiding a cutting of microgrid ruining especially if it is related to a big solar plant, this will make the manager to perform a prior decision regarding the policy of managing all producer equipment of microgrid in technical, economical and environmental point of view. In this thesis, the topic of irradiation solar forecasting have been employed. First the data have been collected from an online platform regarding the meteorological station of Malviya National Institute of Technology from Jaipur region in India, 1922 data have been used including, solar irradiation, wind speed, relative humidity, temperature and sampling time of 15 min. The data have been pre-treated and analyzed by doing a statistical study, removing outliers values based on quarterlies methodology. Furthermore, a matrix of correlation have been calculated to asses the relative humidity, wind speed, temperature regarding the solar radiation variation. The correlation were also calculated to show the impact of having data outlier in statistical behaviour of a data-set. as result of this first part, removing data outliers have increased the correlation between features. Moreover, the results allows to choose the predictors that was: the relative humidity, wind speed , temperature and sampling time.

Supervised machine learning based on regression learning have been used to forecast the solar irradiation, based on the predictors indicated before, the models including regression machine learning based approaches including linear regression models, regression trees, Gaussian process regression models, support vector machines, kernel approximation, ensembles of regression trees, and neural network regression models, have been used to train the model. In order, to assess

the models training accuracy, root mean squared error, mean square error, R-squared and mean absolute error were used. Moreover, the results have been discussed considering with and without data treatment. For both cases, the exponential GPR was giving the best results and showing that even with a short set of historical data, the model can perform a precise predictions. The affect of data perturbation have also been remarked in term of results accuracy, cleaned data outperform the results without processing. The exponential GPR have been exported and tested with a set of 100 data within a sampling time of 15 min for each data including the predictors described previously.

Perspectives

As future work, predicting the value of wind speed is also important for assessing a microgrid system. Envisaging to increase the other models combining optimization algorithm with regression models is important to increase models accuracy.

Bibliography

- [1] S. Automation, “Smart microgrid energy solutions to optimise energy delivery,” 2019. Last accessed 23 February 2022.
- [2] W. Maidana, V. Leite, Â. P. Ferreira, L. Queijo, J. Batista, J. Bonaldo, and E. Goncalves, “Design of a self-sustainable system based on renewable energy sources for a small museum of science dissemination-the house of silk,” in *III Congresso Ibero-Americano de Empreendedorismo, Energia, Ambiente e Tecnologia*, pp. 155–160, Instituto Politécnico de Bragança, 2017.
- [3] L. G. A. Figueiredo, W. Maidana, and V. Leite, “Implementation of a smart microgrid in a small museum: The silk house,” in *Ibero-American Congress of Smart Cities*, pp. 121–134, Springer, 2019.
- [4] C. Marnay, S. Chatzivasileiadis, C. Abbey, R. Iravani, G. Joos, P. Lombardi, P. Mancarella, and J. von Appen, “Microgrid evolution roadmap,” in *2015 international symposium on smart electric distribution systems and technologies (EDST)*, pp. 139–144, IEEE, 2015.
- [5] E. Wood, “Microgrids,” 2020. Last accessed 10 october 2021.
- [6] R. Hacıoğlu, “Prediction of solar radiation based on machine learning methods,” *The journal of cognitive systems*, vol. 2, no. 1, pp. 16–20, 2017.
- [7] C. J. Fluke and C. Jacobs, “Surveying the reach and maturity of machine learning and artificial intelligence in astronomy,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 2, p. e1349, 2020.
- [8] G. F. Cooper, “The computational complexity of probabilistic inference using bayesian belief networks,” *Artificial intelligence*, vol. 42, no. 2-3, pp. 393–405, 1990.
- [9] A. Qazi, H. Fayaz, A. Wadi, R. G. Raj, N. Rahim, and W. A. Khan, “The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review,” *Journal of cleaner production*, vol. 104, pp. 1–12, 2015.

- [10] R. Meenal and A. I. Selvakumar, "Assessment of svm, empirical and ann based solar radiation prediction models with most influencing input parameters," *Renewable Energy*, vol. 121, pp. 324–343, 2018.
- [11] A. K. Yadav, H. Malik, and S. Chandel, "Application of rapid miner in ann based prediction of solar radiation for assessment of solar energy resource potential of 76 sites in northwestern india," *Renewable and Sustainable Energy Reviews*, vol. 52, pp. 1093–1106, 2015.
- [12] A. K. Yadav and S. Chandel, "Solar radiation prediction using artificial neural network techniques: A review," *Renewable and sustainable energy reviews*, vol. 33, pp. 772–781, 2014.
- [13] M. Reza Parsaei, H. Mollashahi, A. Darvishan, M. Mir, and R. Simoes, "A new prediction model of solar radiation based on the neuro-fuzzy model," *International Journal of Ambient Energy*, vol. 41, no. 2, pp. 189–197, 2020.
- [14] S. Al-Dahidi, O. Ayadi, J. Adeeb, M. Alrbai, and B. R. Qawasmeh, "Extreme learning machines for solar photovoltaic power predictions," *Energies*, vol. 11, no. 10, p. 2725, 2018.
- [15] B. Xiao, H. Zhu, S. Zhang, Z. OuYang, T. Wang, and S. Sarvazizi, "Gray-related support vector machine optimization strategy and its implementation in forecasting photovoltaic output power," *International Journal of Photoenergy*, vol. 2022, 2022.
- [16] C. Persson, P. Bacher, T. Shiga, and H. Madsen, "Multi-site solar power forecasting using gradient boosted regression trees," *Solar Energy*, vol. 150, pp. 423–436, 2017.
- [17] H. Sheng, J. Xiao, Y. Cheng, Q. Ni, and S. Wang, "Short-term solar power forecasting based on weighted gaussian process regression," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 1, pp. 300–308, 2017.
- [18] W. Turkenburg and A. Faaij, "Renewable energy technologies (pp. 219–72)," tech. rep., UNDP/UNDESA/WEC: Energy and the challenge of sustainability. World Energy ..., 2000.
- [19] T. M. Razykov, C. S. Ferekides, D. Morel, E. Stefanakos, H. S. Ullal, and H. M. Upadhyaya, "Solar photovoltaic electricity: Current status and future prospects," *Solar energy*, vol. 85, no. 8, pp. 1580–1608, 2011.
- [20] Y. Amoura, *Optimal energy management of a microgrid system*. PhD thesis, 2021.

- [21] N. O. A. Administration, “Weather systems and pattern,” 2008. Last accessed 20 december 2021.
- [22] K. Kaygusuz, “Sustainable development of hydroelectric power,” *Energy sources*, vol. 24, no. 9, pp. 803–815, 2002.
- [23] M. A. S. Calvo, M. V. Llanes, and M. C. Fernández, “Diseño de microrred eléctrica para la comunidad cocodrilo de la isla de la juventud, cuba: Micro grid design for the cocodrilo community on the isla de la juventud, cuba,” *Ingeniería Energética*, vol. 42, no. 3, 2021.
- [24] “Machine learning,” 2017. Last accessed 08 March 2022.
- [25] S. Glen, “Mean squared error: Definition and example,” 2010. Last accessed 08 march 2022.
- [26] S. Glen, “Root mean squared error: Definition and example,” 2010. Last accessed 08 march 2022.
- [27] C. Sammut and G. I. Webb, *Encyclopedia of machine learning*. Springer Science & Business Media, 2011.