Solar Power Probabilistic Forecasting by Using Multiple Linear Regression Analysis in SAS Enterprise Guide®

Mohamed A. Abuella
Deparmtent of Electrical and Computer Engineering
University of North Carolina at Charlotte
Charlotte, USA
Email: mabuella@uncc.edu

December 9, 2014

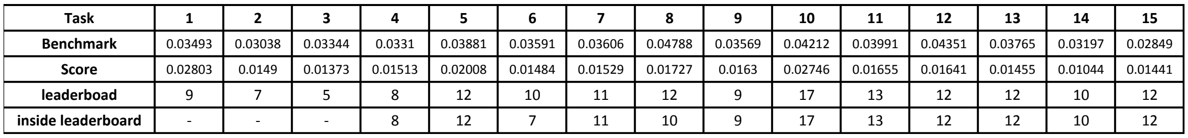
1 Introduction

This is a report about the model that has been used in submission the tasks of probabilistic forecasts for the solar power track in Global Energy Forecasting Competition 2014. This report is based in the basic and main corner stones of the model building stages that led to improve its performance.This multiple linear regression (MLR) analysis model has been built in SAS Enterprise Guide® to generate the forecasts and also MATLAB® is used to assist in data perparation and enhancing the forecasts that are submitted for the competition.

The first section is an introduction for the competition in general and the solar track in particular, in this section the used data description and information about the zones are included. Then the solar forecasting modeling section, where the data preparation steps and the general model building stages are presented. After that, the results section where the results of the model and the evaluation of its performance are represented. Finally the conclusions. In addition, for any further discussions or graphs that lead to loose the organization and the connectivity of the report, these additional materials are replaced in the appendix section.
The illustrative way is adopted rather than the deep theoretical analysis. So the figures, flowcharts, and tables are used to support and clarify basic concepts and ideas.

The following table is for the scores and the ranks of tasks submissions in the solar track of the competition.

Table 1: Scores and Ranks of the Submissions



Task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Benchmark	0.03493	0.03038	0.03344	0.0331	0.03881	0.03591	0.03606	0.04788	0.03569	0.04212	0.03991	0.04351	0.03765	0.03197	0.02849
Score	0.02803	0.0149	0.01373	0.01513	0.02008	0.01484	0.01529	0.01727	0.0163	0.02746	0.01655	0.01641	0.01455	0.01044	0.01441
leaderboad	9	7	5	8	12	10	11	12	9	17	13	12	12	10	12
inside leaderboard	-	-	-	8	12	7	11	10	9	17	13	12	12	10	12

Note: Since the pinball loss function has been built in MATLAB, in last of this report there is Table 4 shows what the scores would have been if the best two accomplished models were used.

2 Global Energy Forecasting Competition 2014 (GEFCom2014)

The Global Energy Forecasting Competition is the second of its time, since the first was in 2012. It was for load and wind point forecasting. While the current competition GEFCom2014 has four tracks: Probabilistic Electric Load Forecasting, Probabilistic Electricity Price Forecasting, Probabilistic Wind Power Forecasting and Probabilistic Solar Power Forecasting to provide state-of-the-art techniques of energy forecasting [1].

2.1 Probabilistc Solar Power Forecasting Objective

In this track of competition, the conants are required to submit the probabilistic forecasts of the solar power in hourly steps through a month of forecasts horizon. It's the track of the competition where the proposed model has been implemented and the data come from.

2.2 Data Descreption

Each submitted task of probabilistic distribution (in quantiles) of the solar power generation is for 3 solar systems (zones).

2.2.1 Zones

The three solar systems are adjacent. These farms are also called zones, they have a relatively low output power capacities. Their installation parameters are:

- Zone1: Altitude (595m) Panel type(Solarfun SF160-24-1M195) Panel number (8) Nominal power (1560W) Panel Orientation (38°Clockwise from North) Panel Tilt (36°).
- Zone2: Altitude(602m) Panel type(Suntech STP190S-24/Ad+) Panel number (26) Nominal power (4940W) Panel Orientation (327°Clockwise from North) Panel Tilt (35°).
- Zone3: Altitude(951m) Panel type(Suntech STP200-18/ud) Panel number (20) Nominal power (4000W) Panel Orientation (31°Clockwise from North) Panel Tilt (21°)

2.2.2 Weather variables

The target variable is the solar power. There are 12 NWP independent variables are from European Centre for Medium-Range Weather Forecasts (ECMWF). The ECMWF-NWP output to be used as below:

- 078.128 Total column liquid water of cloud (tclw) - (kg/m^2).
- 079.128 Total column ice water of cloud (tciw) - (kg/m^2)
- 134.128 surface pressure (SP) - (Pa).
- 157.128 Relative humidity at 1000 mbar (r) - (%).
- 164.128 total cloud cover (TCC) - (0-1)
- 165.128 10 metre U wind component (10u) - (m/s).
- 166.128 10 metre V wind component (10V) - (m/s).
- 167.128 2 metre temperature (2T°) - (K)
- 169.128 surface solar rad down (SSRD) - (J/m^2) - Accumulated field.
- 175.128 surface thermal rad down (STRD) - (J/m^2)- Accumulated field.
- 178.128 top net solar rad (TSR) - (J/m^2) -Accumulated field.
- 228.128 total precipitation (TP) - (m) - Accumulated field.

2.3 Evaluation Creteria

Pin ball loss function is used to evaluate the accuracy of probabilistic forecasts. It's a piecewise linear function is often used to evaluate the accuracy of quintile forecasts [2].

$$L_q(\hat{Y}, Y) = \begin{cases} q(\hat{Y} - Y), & \text{if } Y \leq \hat{Y} \\ (1-q)(Y - \hat{Y}), & \text{if } Y > \hat{Y}. \end{cases} \tag{1}$$

Where $L_q(\hat{Y}, Y)$ is the loss function to the probabilistic forecasts for each hour. $\hat{Y}$ is the forecasted value at the certain q quantile of the probabilistic solar power forecasts and Y is the observed value of the solar power. The quantile q is also called the percentile and it has discrete values [0.01 - 0.99]. $\hat{Y}$ and Y are normalized values of the nominal power capacity of each zone. The average of loss function $L_q(\hat{Y}, Y)$ for all forecasting hours should be minimized to yield more accurate forecasts. Therefore, the average loss function value is adopted as a score to evaluate the model performance.

3 Solar Forecasting Modeling

The basic skeleton of the methodology and theoretical background of building the proposed solar power forecasting model in this report is adopted from reference [3], where the multiple linear regression (MLR) analysis is deployed for short-term load forecasting. SAS Enterprise Guide® is used as the environment where the forecasting model is built in. SAS is a statistical analysis system software is suitable for the proposed statistical forecasting model. The flowchart of the solar forecasting model building steps is shown in Figure. 1.

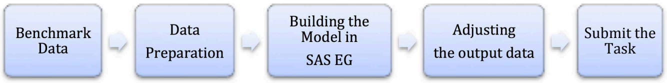


Figure 1: Flowchart diagram of the solar forecasting modeling

3.1 Data Preparation

It always a good idea to get the analysis of the historical data before setting up the forecasting model. The available historical data contains the solar power and all 12 weather variables for each zone. The data preparation is an important step for treating the data to be ready for the analysis and modeling.
The various steps of the data preparation are shown in Figure. 2. Figure. 3 shows the box plot of the distribution of the historical data (i.e.2012) of solar power for zone1. The other two zones are not different a lot since they are adjacent solar systems. Note, the order of months as it is provided from the competition benchmark data and appears in the box plot does not necessary mean the same order of actual year months.

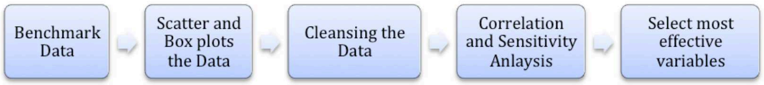

Benchmark Data
Scatter and Box plots the Data
Cleansing the Data
Correlation and Sensitivity Anlaysis
Select most effective variables

Figure 2: Flowchart diagram of data preparation

The scatter plot is also useful to get sense of the relationships between the explanatory variables (the weather variables) and response variable (the solar power). Figure. 4 presents the advantage of plotting the data since the scatter plots for the observed power with respect to the solar irradiance as it is also

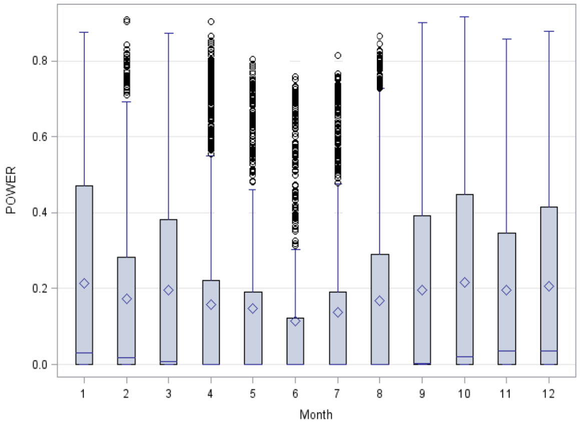
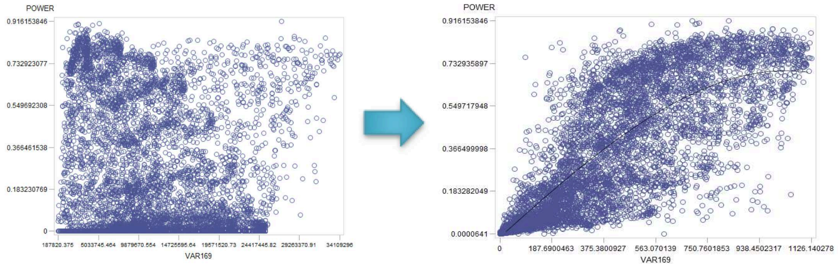


Figure 3: Box plot of the distribution of observed solar power

called surface solar radiation down (SSRD) for zone1. The scatter plot on the left in Figure. 4 is for the solar irradiance SSRD169.128 as it is provided from the NWP which is in accumulated values (J/m^2), while the plot on the right for the average values of solar irradiance SSRD (W/m^2). The relationships between the variables on the right plot is more obvious and it can tell it is a positive relationship with relatively high positive correlation coefficient. The last four given weather variables (i.e. solar and thermal radiations besides the precipitation) are in accumulated field values not average values. They are increasing in every hour until the end of the day and then start again in accumulation [4]. For getting the average values of these data by applying the formula in equation 2.



169.128 (SSRD) : J/m^2 **Accumulated field (for all day hours)**

169.128 (SSRD) w/m^2 **Average Value (for each hour)**

Figure 4: Scatter plot of the observed solar power vs. solar irradiance

$$Avg(t) = \frac{Acc(t+1) - Acc(t)}{3600} \tag{2}$$

Since t is the time in hour steps, Avg and Acc are the average and accumulated values of the data respectively.

For comparison of the three zones, from Figure 5, zone3 has a strongest positive linear relationship between the solar irradiance and power. The left scatter plot is for zone1.

From the scatter plots, the outliers don't change the general data trends. The majority of the extreme points in the observed data occur at the sun rise and set periods which are the unstable duration of solar panel. By experiment the data cleansing is conducted and led to a tiny improvement in forecasts. This doesn't mean to underestimate the data cleansing stage in data preparation before the modeling since sometimes the outliers could be generated from data entry issues. See the appendix for methodology that have been used in data cleansing.

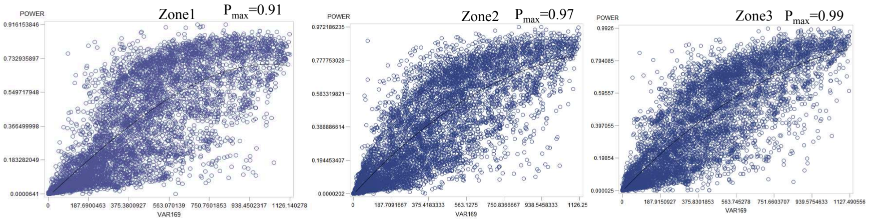


Figure 5: Scatter plot of the solar irradiance and power for all zones

3.2 The Model Building

The main steps of building the forecasting model is shown in Figure 6. ANOVA is the analysis of variance and it's available as a statistical tool in SAS where the multiple linear regression analysis is running.

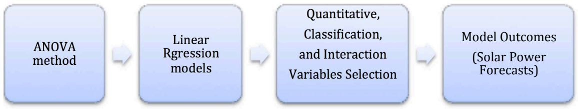

ANOVA method
Linear Rgression models
Quantitative, Classification, and Interaction Variables Selection
Model Outcomes (Solar Power Forecasts)

Figure 6: Flowchart diagram of the model building

The explanatory variables selection for forecasting the solar power as independent or response variable by plotting is cumbersome since there is twelve explanatory variables to choose form, see figures that are shown in the appendix. So the correlation and sensitivity analysis is carried out for the historical data in SAS Rapid Predictive Modeler to investigate the most effective variables. Figure 7 is the outcomes of the Rapid Predictive Modeler in SAS.

It's obvious the solar irradiance, in addition to the time (hours), the surface irradiance (VAR169) and net top solar irradiance (VAR178) and their second order polynomial or quadratic terms have the highest correlation with solar power. The relative humidity (VAR157) has some impact on solar power since the water vapor particles by chance work as small mirrors and increasing the reflection level of the solar irradiance and hence the generated solar power. The temperature at 2m variable (VAR167) is also has an noticeable impact compared to other less effective variables. It is reasonable since the solar panel characteristics are affected by the ambient temperature. In fact, the temperature plus to the solar irradiance are chosen in physical-approach models to forecast the power from PV systems [5].

The multiple linear regression MLR model can be represented as in equation 3.

$$Y = \beta_o + \beta_1 X_1 + ... + \beta_k X_k + \epsilon \tag{3}$$

Where Y is the response variable and X_k is the k explanatory variable. βs are the parameters and ϵ is the error or the variations in Y. The response variable is the solar power, and the explanatory variables are selected from the most effective weather variables [3].

In multiple linear regression analysis model, the ability of using the interactions between the variables adding more power to the model to produce more accurate forecasts. In addition, MLR analysis model includes the probability of the forecasts as a prediction interval. The prediction interval includes the variance of the error as well as the variance of the parameter estimates.

Since the sun azimuth angle changes periodically through the day hours and the sun elevation changes along the seasons [6]. Therefore, the hours and month days and months are considered as classification variables that the chosen explanatory variables interact with.

It's worth to mention, that the variables selection is done based on each separated variable alone, but when the variables interact with each other in the MLR analysis model they could loose some correlation

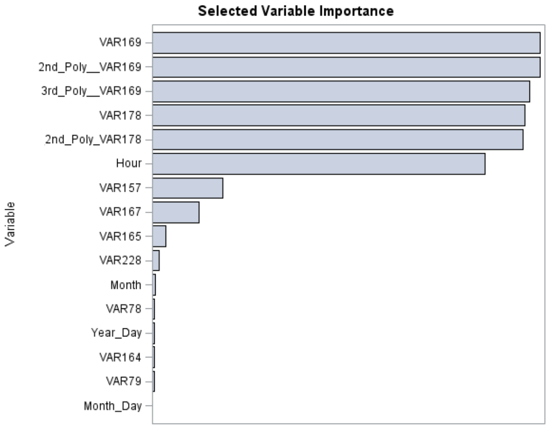


Figure 7: The order of most effective variable for solar power forecast

power. Therefore, in the total mix of selected variables, the best candidate model is also needed. After the variables selection and their interactions are investigated, the selection of the candidate model of the most efficient performance is found manually by carrying out the three main steps of building the model, training, validation, and testing. This is carried out by using Pinball loss function to eliminate the less efficient varibales every time that they are discovered since these variable don't lead to improve the performance of the model. Thus, eventually the candidate model was found with the most efficient variables that altogether work to enhance the forecasts of the model.

Figure. 8 shows a part of the process flow of the solar forecasting model in SAS Enterprise Guide project window.

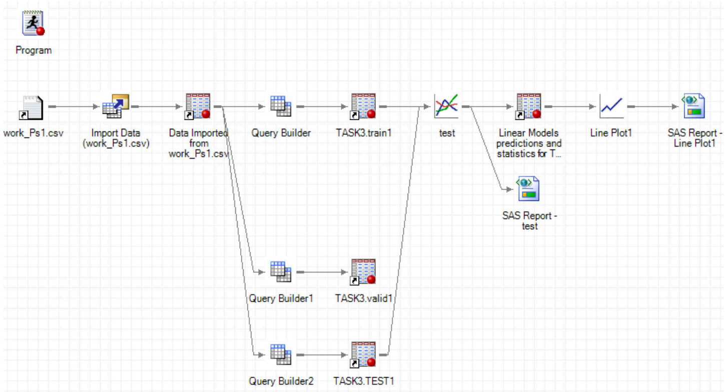


Figure 8: Diagram shows a part of the process flow of the model in SAS-EG

4 The Model Results and Evaluation

The sample probabilistic forecasts of the solar power for each hour and for all three zones or solar systems are as in Table 2. The forecasts in median column ($q = 0.50$) is the point forecast which should be as close as possible to the actual observed power.

Since after submission the forecasts for a task, then in the following task the actual data for these forecasts

Table 2: Sample of probabilistic forecasts for some hours of zone1

ZONE ID	TIME	0.01	0.02	0.03	.....	0.49	0.5	0.51	.....	0.97	0.98	0.99
1	20130401 01:00	0.223363	0.227528	0.231693	.....	0.535853	0.555	0.55722	.....	0.621648	0.625593	0.629539
1	20130401 02:00	0.074106	0.075488	0.07687	.....	0.423273	0.438397	0.440151	.....	0.206246	0.207555	0.208864
1	20130401 03:00	0.057057	0.058121	0.059185	.....	0.140431	0.145449	0.146031	.....	0.158798	0.159806	0.160814
1	20130401 04:00	0.029166	0.029709	0.030253	.....	0.108124	0.111987	0.112435	.....	0.081171	0.081687	0.082202

is provided, the plots can be used as a validation of the model performance. As it is shown in Figure. 9 the actual and the point (median) forecasts are plotted with residual plot. The residuals plot has positive and negative values these mean not over-forecasting nor under-forecasts impacts, so the model is not biased. Since the forecasts is probabilistic, the box plots for the distribution of the actual and forecasted power is also shown for each day hour.

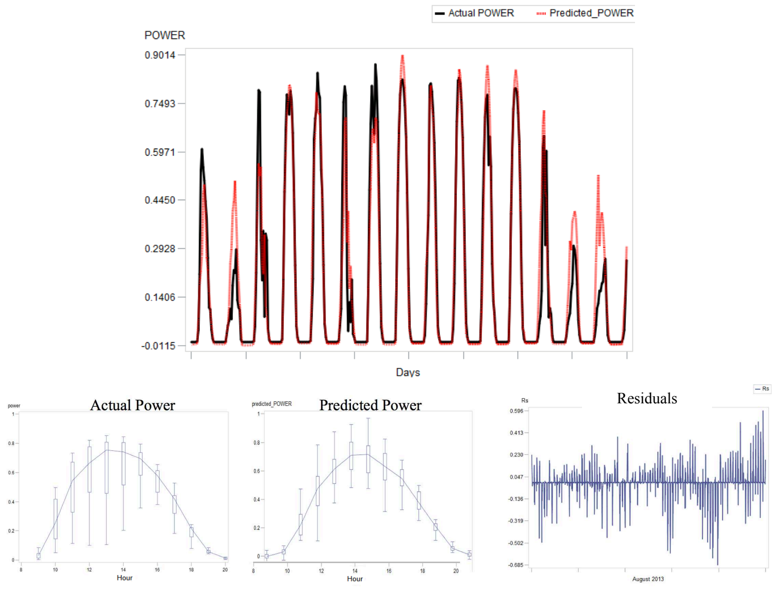


Figure 9: The results plots for solar power forecasts of the model

The pinball is also built and used for evaluate the mode as well as the root mean square error (RMSE) between the point power forecasts and actual solar power. Table 3 presents the pinball loss function value and RMSE for each zone and for the entire aggregated system of the three zones.

Table 3: September 2013 Forecasts evaluation

Zone	Pinball	RMSE
1	0.0130	0.0725
2	0.0140	0.0741
3	0.0141	0.0742
Aggregated	0.0137	0.0736

The line plot is shown in Figure. 10 for the actual solar power and its corresponding forecast for the median $q = 0.5$ and both the lower and upper limits of prediction interval if the the prediction interval is set up to 95% this makes the prediction interval contains 95% of the observed solar power for each hour. The other quantiles forecasts ($q = 0.02$ to 0.49 and $q = 0.51$ to 0.98) for each hour are distributed linearly between the upper and lower limits ($q = 0.01$ and $q = 0.99$) of the prediction interval.

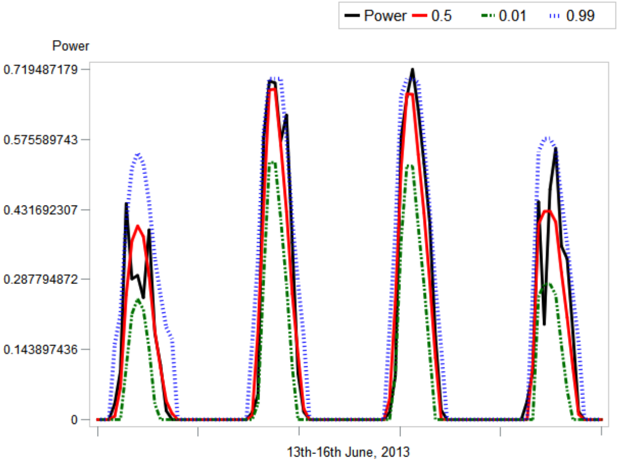


Figure 10: The line plots for actual solar power and the forecasts where $q = 0.01, 0.5, 0.99$

Meanwhile the scatter plots of the same quantiles for zone1 is shown in Figure. 11. It's clear that the trend lines are closed in the point forecasts when $q = 0.50$ and actual power. From line and scatter plots, it's obvious that the probabilistic forecasts produced by the model covers the actual solar power.

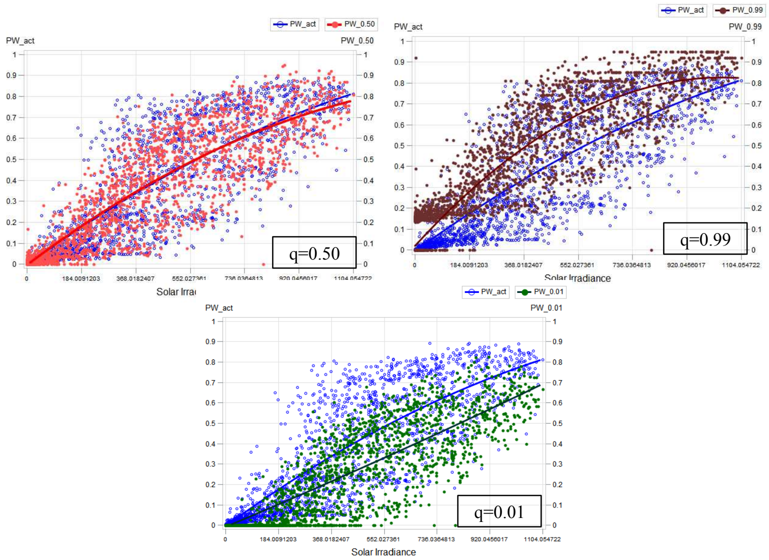


Figure 11: The scatter plots for actual solar power and the forecasts where $q = 0.01, 0.5, 0.99$

Another useful evaluation way of the model is when the scatter plot of the point forecasts ($q = 0.50$) are drawn with respect to the observed power and then finding the correlation. This is represented by scatter plot in Figure. 12, where Pearson correlation coefficient is found 0.9414.

Table 4 shows the pinball loss function (scores) if the best two models were used for submissions of

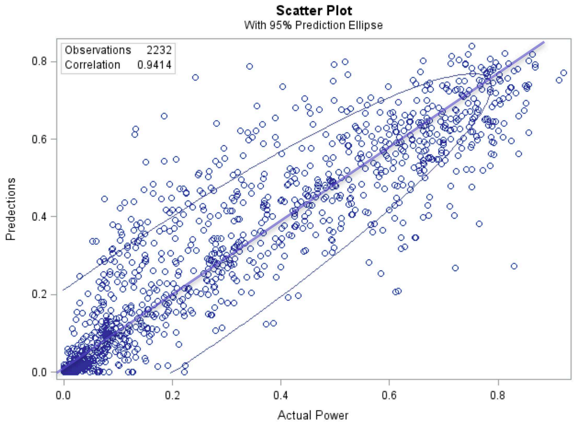


Figure 12: Scatter plot of the point forecasts vs the actual solar power

all monthly forecasts tasks. This experiment was conducted by assistance of MATLAB to calculate the loss function between the solar power forecasts and the actual power, these monthly actual solar power were obtained as historical data from the following task after the previous task had been submitted. I was thinking that one model is fit for cloudy months and the other for sunny months, that is why the months are used as identification of columns rather than the task numbers. However, the results show that the scores of both models are not so different and they are not correlated directly to the months' status whether they are sunny or cloudy months. The code of the candidate model ($model1$) that uses the most efficient variables and hence less degree of freedom is shown in the appendix.

Table 4: The monthly forecasts scores if best Models were used

Month	April	May	June	July	August	September	October	November	December	January	February	March	April	May	June
Benchmark	0.03493	0.03038	0.03344	0.0331	0.03881	0.03591	0.03606	0.04788	0.03569	0.04212	0.03991	0.04351	0.03765	0.03197	0.02849
MODEL1	0.0147	0.0129	0.0131	0.0153	0.019	0.014	0.0151	0.0173	0.0171	0.0158	0.0157	0.0163	0.0138	0.01061	
MODLE2	0.0149	0.013	0.0131	0.0156	0.0191	0.0136	0.0149	0.0177	0.0169	0.0163	0.0163	0.0161	0.0146	0.01044	

Note: For the last task, the pinball loss function couldn't be calculated because the actual power data is not available.

5 Conclusion

The adopted multiple linear regression analysis model as a statistical approach model performed well for solar power forecasts. Since no deep comparison study has been conducted with other models such as artificial intelligence and physical approach models, the model might not outperform other models. If the forecasting hours were with clear sky, the model's performance would be better for near forecasting horizon than farther horizon but this is affected by cloudy hours and so the entire performance of the model. Plotting the data, correlation and sensitivity analysis between the variables, as well as data cleansing of outliers if there any are essential data preparation steps before building the forecasting model. For additional historical data the model performance will be improved. By experiment, found the quadratic term of the solar irradiance is doing better job than the cubic term and the natural logarithm term. Since the hours selected as a classification variables, deleting the night hours from data reduces the model's performance.

Acknowledgment

I would like to express my grateful thanks to Dr.Tao Hong at University of North Carolina at Charlotte and the General Chair of Global Energy Forecasting Competition 2014 for his cooperation, valuable feedbacks, and enlightening blogs. Additional thanks to Dr.Badrul Chowdhury for his insightful discussions about energy forecasts in general. Last but not least, the useful talks and chats with the classmates, whether they are at campus or online, they are greatly appreciated as well.

References

[1] "Global Energy Forecasting Competition 2014, Probabilistic solar power forecasting." [Online]. Available: http://www.crowdanalytix.com/contests/global-energy-forecasting-competition-2014.

[2] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets.* Springer, 2014.

[3] T. Hong, *Short term electric load forecasting.* North Carolina State University, 2010.

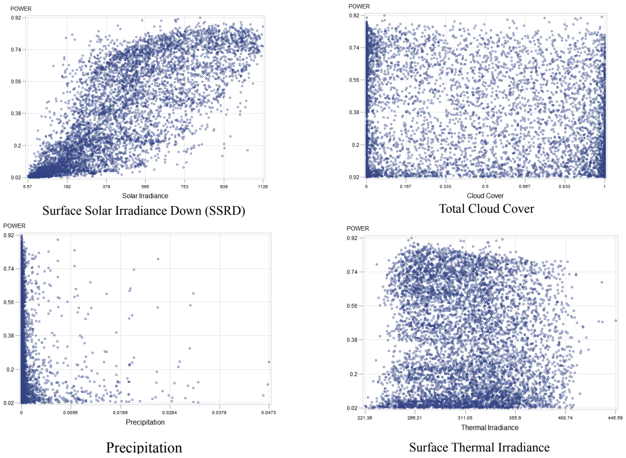
[4] "Many fields have seconds in their units e.g. radiation fields. How can instantaneous values be calculated?." [Online]. Available: http://www.ecmwf.int/en/many-fields-have-seconds-their-units-eg-precipitation-and-radiation-fields-how-can-instantaneous.

[5] E. Lorenz, T. Scheidsteger, J. Hurka, D. Heinemann, and C. Kurz, "Regional pv power prediction for improved grid integration," *Progress in Photovoltaics: Research and Applications*, vol. 19, no. 7, pp. 757–771, 2011.

[6] P. Bacher, H. Madsen, and H. A. Nielsen, "Online short-term solar power forecasting," *Solar Energy*, vol. 83, no. 10, pp. 1772–1783, 2009.

6 Appendix

The appendix contains some illustrative figures, such as graphs and tables since sometimes the picture is worth of thousand words. Figure. 13 shows scatter plots of some variables that are used as beginning investigations if there are any relationships between the solar power and these variables.



Surface Solar Irradiance Down (SSRD)

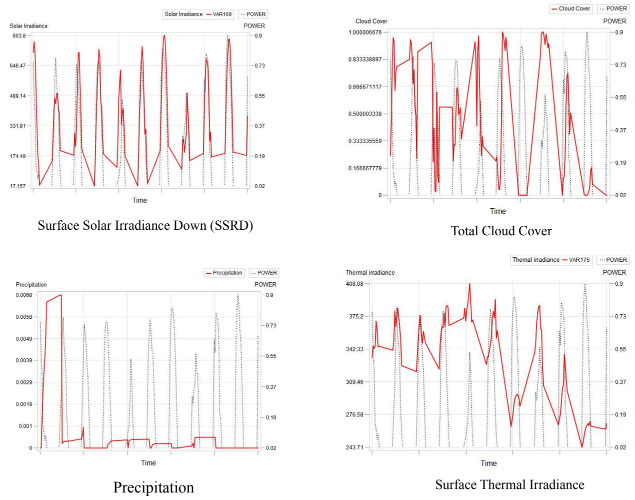
Total Cloud Cover

Precipitation

Surface Thermal Irradiance

Figure 13: Scatter plots of some variables

Figure. 14 shows line plots of the same variables that used in previous scatter plot but now with line plots with two different scales (right for solar power and left for the target variable). The polynomials degree has been investigated for the solar irradiance variable to check out which degree is suitable for the term that is used in the model to increase the accuracy of the forecasts. Figure. 15 shows three degrees of the polynomials, the only quadratic and cubic polynomial are used in interaction terms of the MLR model. In Task 9, the forecasts were for December 2013, there was an outlier in the top net solar irradiance. Since this variable is used in the model this outlier led to extreme forecasts as shown in Figure 16. Some data cleansing has been done by assistance of MATLAB to bring the extreme points to the fitted curve which is in the middle of the data. The adopted strategy of data cleansing is shown if Figure. 17. However, the data cleansing led to a tiny improvement in the forecasts.



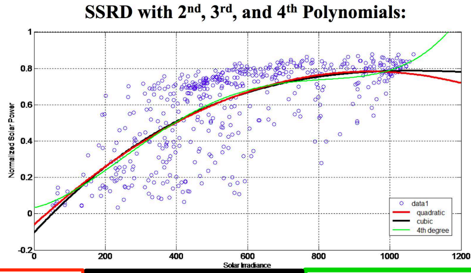
Surface Solar Irradiance Down (SSRD)

Total Cloud Cover

Precipitation

Surface Thermal Irradiance

Figure 14: Line plots of the same variables that are shown in Figure. 13



Quadratic:
y = p1*x^2 + p2*x + p3

Coefficients:
p1 = -9.4913e-07
p2 = 0.0017901
p3 = -0.061731
Norm of residuals =
3.6151

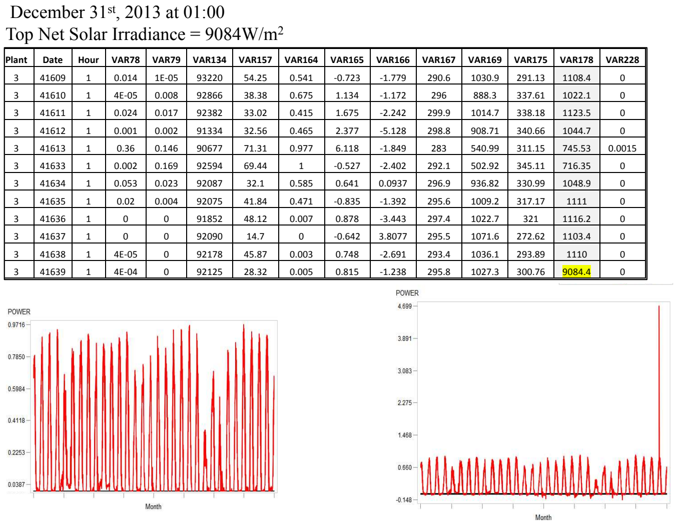
Cubic:
y = p1*x^3 + p2*x^2 +p3*x + p4

Coefficients:
p1 = 3.848e-10
p2 = -1.6087e-06
p3 = 0.0021135
p4 = -0.10454
Norm of residuals =
3.6122

4th degree Polynomial:
y = p1*x^4 + p2*x^3 + p3*x^2 + p4*x + p5

Coefficients:
p1 = 3.5244e-12
p2 = -7.5065e-09
p3 = 4.3011e-06
p4 = 0.00043816
p5 = 0.033125
Norm of residuals =
3.5957

Figure 15: Curve fitting plots in MATLAB for polynomials of solar irradiance variable, SSRD

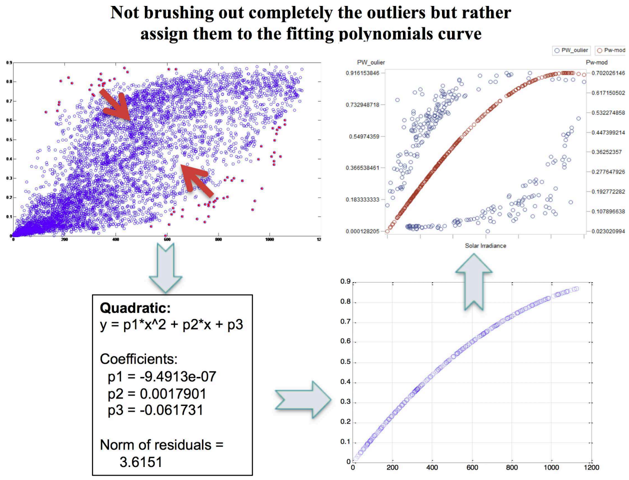


December 31st, 2013 at 01:00

Top Net Solar Irradiance = 9084W/m2

Plant	Date	Hour	VAR78	VAR79	VAR134	VAR157	VAR164	VAR165	VAR166	VAR167	VAR169	VAR175	VAR178	VAR228
3	41609	1	0.014	1E-05	93220	54.25	0.541	-0.723	-1.779	290.6	1030.9	291.13	1108.4	0
3	41610	1	4E-05	0.008	92866	38.38	0.675	1.134	-1.172	296	888.3	337.61	1022.1	0
3	41611	1	0.024	0.017	92382	33.02	0.415	1.675	-2.242	299.9	1014.7	338.18	1123.5	0
3	41612	1	0.001	0.002	91334	32.56	0.465	2.377	-5.128	298.8	908.71	340.66	1044.7	0
3	41613	1	0.36	0.146	90677	71.31	0.977	6.118	-1.849	283	540.99	311.15	745.53	0.0015
3	41633	1	0.002	0.169	92594	69.44	1	-0.527	-2.402	292.1	502.92	345.11	716.35	0
3	41634	1	0.053	0.023	92087	32.1	0.585	0.641	0.0937	296.9	936.82	330.99	1048.9	0
3	41635	1	0.02	0.004	92075	41.84	0.471	-0.835	-1.392	295.6	1009.2	317.17	1111	0
3	41636	1	0	0	91852	48.12	0.007	0.878	-3.443	297.4	1022.7	321	1116.2	0
3	41637	1	0	0	92090	14.7	0	-0.642	3.8077	295.5	1071.6	272.62	1103.4	0
3	41638	1	4E-05	0	92178	45.87	0.003	0.748	-2.691	293.4	1036.1	293.89	1110	0
3	41639	1	4E-04	0	92125	28.32	0.005	0.815	-1.238	295.8	1027.3	300.76	9084.4	0

Figure 16: The outlier in top net solar irradiance variable and its consequences



Not brushing out completely the outliers but rather assign them to the fitting polynomials curve

Quadratic:
y = p1*x^2 + p2*x + p3

Coefficients:
p1 = -9.4913e-07
p2 = 0.0017901
p3 = -0.061731

Norm of residuals =
3.6151

Figure 17: Data cleansing for solar irradiance variable by using curve fitting in MATLAB

The candidate model is as following:
CLASS = Month ”Month Day”n Hour; MODEL POWER = VAR167 VAR169 VAR175 VAR228 ”2VAR169”n Month Hour VAR79 VAR157 ”3VAR169”n VAR169*Hour VAR169*Month ”Month Day”n*VAR169 ”2VAR169”n*Month ”2VAR169”n*Hour VAR167*VAR169*Hour VAR169*VAR175*Hour VAR169*VAR157*Hour VAR169*VAR228*Hour VAR169*VAR79*Hour ”Month Day”n”2VAR169”n*”Month Day”n Trend ”4VAR169”n ”5VAR169”n ”Year Day”n”4VAR169”n*Month”4VAR169”n*Hour ”4VAR169”n*”Month Day”n VAR78 VAR134 VAR164 VAR165 VAR166 VAR178 VAR178*Hour VAR178*Month ”Month Day”n*VAR178 VAR79*Hour VAR228*Hour VAR157*Hour VAR175*Hour VAR167*Hour VAR134*Hour VAR78*Hour VAR164*Hour VAR165*Hour VAR166*Hour”3VAR169”n*Hour”3VAR169”n*Month”3VAR169”n*” MonthDay”n”5VAR169”n*Hour”5VAR169”n*Month