IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-10, Oct 2018 E-ISSN: 2347-2693

Performance Comparison of Forecasting on Solar Plant Generation Data

Sukhpal Kaur[1*], Madhuchanda Rakshit[1]

[1]Research Scholar, Guru Kashi University, Talwandi Sabo, Bathinda, Punjab, India
[1]Assistant Professor, Applied Science Department, Guru Kashi University , Talwandi Sabo, Bathinda, Punjab, India

[*]Corresponding Author: sukhpakaur.sond@gmail.com Tel.: +91-9256310300

Available online at: www.ijcseonline.org

Accepted: 20/Oct/2018, Published: 31/Oct/2018

***Abstract*—** The purpose of this paper is to compare the forecasting performances of daily generation of a solar plant by utilizing the autoregressive time series models. As the demand for energy is increasing frequently all over the world, the proper integration of solar energy and its accurate predictions become necessary for our society for better planning and distribution of energy. In this study, we compare our solar energy time series data with Autoregressive Integrated Moving Average with Explanatory Variable (ARIMAX) and Vector Autoregressive (VAR) time series models for analyzing our solar plant data separately and at last conclusion is made on the better performance of these two methods. Moreover for VAR model effects of various variables are tested for maximum production of solar power. For evaluating the accuracy performance of our forecasted data, we use Mean Absolute Error (MAE), Mean Absolute Scaled Error (MASE) and Root Mean Square Error (RMSE) measurements.

***Keywords*—** ARIMAX, MAE, MASE, RMSE, Solar Plant Generation, VAR

I. INTRODUCTION

The fastest growing population of the world increases the challenges for the industries to enhance their energy requirements. Solar energy is the most important alternative resource of the world and has a large potential of green energy. Keeping in view the limitation of the conventional resources, the industries all over the world has to think about the alternative source of energy. Recently most of the countries of the world have been emphasizing on the development of renewable energy resources. Renewable energy plays an important role in this regard as this energy is abundant and free of cost from nature.

Solar power in India is a developing industry. The installed capacity of our country reached 20 GW in February 2018. The increasing demand for energy is one of the biggest reasons behind the integration of solar energy into the electric grids. To ensure the efficient use of energy systems, it becomes important to forecast information reliable.

Solar plant generation forecasting is the fundamental basis of managing existing and newly constructed power systems. Without having accurate predictions for the generated power, serious implications such as appropriate operational practice and inadequate energy transaction are inevitable. Mohamed Abuella and Badrul Chowdhury [15] were recently studied the solar power forecasting using support vector regression. Jabar H. Yousif and Hussein A Kazem [11] in their paper had addressed about Modeling of daily solar energy system prediction using soft computing methods.

Forecasting is a necessary aid to planning and it is use for historical data to determine the direction of future trends. There are various models used to extrapolate the solar generation time series data into the future data. Out of various models available in time series in this paper, we use the Autoregressive Integrated Moving Average with explanatory variable (ARIMAX) model and Vector Autoregressive (VAR) model. Many researchers had discussed about ARIMAX and VAR modelling and forecasting of future data. Paul S. P. Cowperwait and Andrew V. Metcalfe in their book “Introductory Time Series with R”[16] and in the book on “Time Series Analysis and Its Applications With R Examples” by Robert H. Shumway and David S. Stoffer [19] gave comprehensive study about the autoregressive time series models which is very useful for all time researchers.

The standard ARIMA (Autoregressive Integrated Moving Average) model allows making forecasts based only on the past values of the forecast variable. The ARIMAX model is an extended version of the ARIMA model. It includes also other independent variables. The model is also referred to as the vector ARIMA or the dynamic regression model. The ARIMAX model is similar to a multivariate regression model, but allows taking advantage of autocorrelation that may be present in the residuals of the regression to improve the accuracy of a forecast. VAR models are used for multivariate time series. The structure is that each variable is a linear function of past lags of itself and past lags of the other variables.

In this research paper, we compare the performance of time series models based on error measures and forecast the future generation of solar power plant. The ARIMAX and VAR models were applied to the daily solar power generation of a plant and analyze the data and then fitted ARIMAX and VAR model using R software. Later on compare them and forecast the future trend of best fitted model. The present paper is organized as follows- Introduction is provided in Section I of this paper. In section II recent related work is provided as compared to our work, in section III our source of data and its details are provided, Methodology of this paper is also provided in section IV, Section V is deals with the elaborated results of our two models and their discussions and finally in section VI conclusion and future scope of this paper is discussed.

II. RELATED WORK

The development and uses of solar energy at large scale is not the only reasonable methods of utilization of solar power, but also very much effective to improve the crisis of resources economically.Harendra Kumar Yadav, Vijay Kumar and Vinay Kumar [10] were recently research about the potential of solar energy in India and given a review on it.

D.K. Chaturvedi and Isha [6] had contributed a review on solar power forecasting. Jai Singh Arya, Aadesh Kumar Arya and Sanjev Aggarwal [12] were working on recent trends in solar energy generation and concluded that solar power plant may be installed in such a way so these may work in unison with hydro and other methods of generation to enhance the clean and green energy.

ARIMAX models have been already applied to forecast various fields, such as vehicular traffic flow, paddy production, sales of food retail industry, oil export, tertiary industry etc. [2,3,14,22]. Javed Iqbal [13] had studied and compared the techniques of ARIMA, VAR, ECM (Error correction Model) and ARCH /GARCH models, and S.M. Husnain Bokhari and Mete Feridun [20] had also provided a comparative analysis between ARIMA and VAR models.Wiwik Anggraeni, Retno Aulia Vinarti, Yuni Dwi Kurniawati [24] studied on performance comparison between Arima and Arimax method in Muslim kids clothes demand forecasting in 2015. Apart from above research work there are also many other researchers who contributed on ARIMAX models, few of them are Herui Cui and Xu Peng [8], A. Jalalkamli, M. Moradi and N. Moradi [1], Christopher Bennett, Rodney A. Stewart and Junwei [4] and Renny Elfira Wulansari, Setiawan and Suhartono [18] had ARIMAX model compared with hybrid models.Wiwik Anggraeni, Kuntoro Boga Andri, Sumaryanto and Faizal Mahananto [23] was recently studied the performance of ARIMAX and VAR model in forecasting, strategic commodity price in Indonesia , Ghulam Ali [9] and El Mostafa Bentour [7] also studied on cointegration VAR , VECM and ARIMAX for water quality and ranking of VAR and structural models in forecasting in 2015.

III. DATA DESCRIPTION

The data consider in this study include daily power generation of 1 MW SPV (Solar Photo- Voltaic) Power Plant of Vill. Phulokhari Distt. Bathinda, Punjab in India. Data was sourced from the Punjab Energy Development Agencies (PEDA) for the time period from January, 2017 to February, 2018. A brief desceiption of the variables used for analysis is as given in table below:

Table 1: Variable Dercription

Variable	Defination
Daily Average Solar Radiation	Solar radiance is the power per unit area received from the sun in the form of electromagnetic radiation in the wave length range of the measuring instrument. The solar irradiance integrated over time is called solar radiation, insolation, or solar exposure. However, insolation is often used interchangeably with radiance in practice
Average Module Temprature	Solar panel temperature is one of the important factors that affects how much electricity your panels will produce.
Average Air Temprature	The temperature of air is a measure of the average thermal energy of the molecules in the air the higher temperature, the higher energy of the molecules.
Solar Power Generation	Solar energy generation is one of fastest growing and most promising renewable energy sources of power generation worldwide

In this study we used the two autoregressive time series model for analysis our data and compare the performance of these model based on error measures for forecasting the solar power generation.

IV. METHODOLOGY

Time series methods are one of the most commonly used statistical techniques for forecasting. Time series is a unit of observation by the data in time order. Time series models are very useful models when you have serially correlated data. We used two major methods ARIMAX and VAR models for solar forecasting; classical statistical techniques, computational intelligent methods.

3.1 Autoregressive Integrated Moving Average with Explainatory Variable (ARIMAX)

The Box – Jenkins method is the way to find the proper model for estimating parameters the time series values by using Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). The model is composed of three major parts, which are Autoregressive Model (AR(p)), Integrated Process(I(d)) and Moving Average Model (MA(q)).The autoregressive integrated moving average (ARIMA) model is used for non stationary time series . A stationary ARMA (p, q) process with the dth difference of the time series develops an ARIMA (p, d, q) model. Then ARIMA (p, d, q) is represented by

$$\overline{x} = \sum_{i=1}^{p} \phi_i B^d x_{t-i} + \sum_{j=0}^{q} \theta_j w_{t-j}$$

ARIMAX is the method that was adopted by including X in ARIMA, which is the leading variable (or Structural Variable or Exogenous Variable) to improve the forecasting accuracy. It is more applicable to time series with sudden changes in trends. An ARIMA(p,d,q) process including m past values of an exogenous variable e_t develops an ARIMAX process of order (p,d,q,m) model is represented by above equation :

$$\overline{x_t} = \sum_{i=1}^{p} \phi_i B^d x_{t-i} + \sum_{j=0}^{q} \theta_j w_{t-j} + \sum_{k=1}^{m} \lambda_k e_{t-k}$$

Where w_t is the white noise. ϕ_i, θ_j and λ_k are the coefficients of the autoregressive , moving average and exogenous inputs , respectively.

3.2 Vector Autoregressive (VAR)

The Vector Autoregressive (VAR) model are a vector of two or more interrelated variable in which each vector component is a function of its own past values and the past values of the other components of the vector to a finite order p (lags). The important step in VAR modelling is the determination of the order of such lags. The Vector Autoregressive model characterizes linear dependence between two or more time series . VAR model uses multiple variables to generalize the univariate autoregressive model (AR model) . A k-dimensional VAR model of order L which is given as below:

$$\overline{x_t} = v + \sum_{i=1}^{L} A_i x_{t-i} + w_t = v + A_1 x_{t-1} + ... + A_L x_{t-L} + w_t$$

Where x_t and v are k$\times$ 1 vectors of variables and constants, respectively. L is the maximum lag in the VAR model, A_i is a k$\times$ k matrix of lag order parameters, and w_t =(w_{1t} ,….., w_{kt}) is the vector of white noise.

3.3 Performance Measures

Time series forecast performance measures gives us a summary of the skill and capability of the forecast model which made the forecasting. We used performance measures which are MAE, MASE and RMSE in our analysis. Higher values of errors correspond to less forecast accuracies.

Mean Absolute Error (MAE)

The mean absolute error is calculated by:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |\tilde{x} - x_i|$$

Where x represents the observed value, $\tilde{x}$ is the predicted value (forecast) and n is the total number of samples.

Mean Absolute Scaled Error (MASE)

The mean absolute scaled error is calculated by :

$$MASE = \frac{\sum_{i=1}^{n} |\tilde{x}_i - x_i|}{\frac{n}{n-1} \sum_{i=2}^{n} |\tilde{x}_i - x_{i-1}|}$$

Similarly as above mentioned.

Root Mean Square Error (RMSE)

The root mean square error is given by calculating the square root of mean square error as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\tilde{x}_i - x_i)^2}$$

Similarly as above mentioned.

V. Result and Discussion

This section contains analysis of generation of solar power plant based on ARIMAX and VAR time series model and finally the evaluation. Intially our analysis is conducted by displaying the summary statistics of the series involved as well as the corresponding time series plot of our time series data with different variables of daily generation of solar plant. The plot of graph and summary table of our daily data which is given as follows:

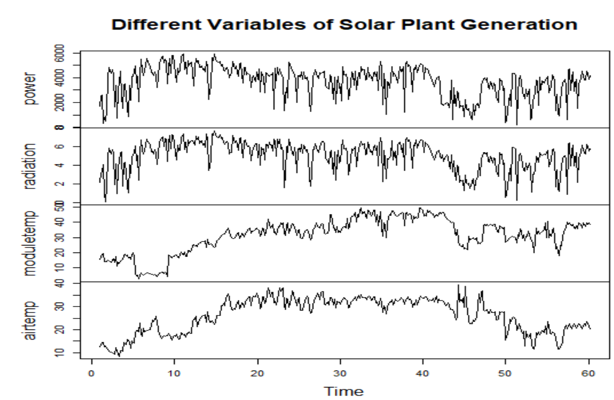

Figure 1. Graph of solar plant generation

Table 2: Summary of Data

	Power (kwh)	Radiation (kwh/ m^2)	Moduletemp (C^0)	Airtemp (C^0)
Minimum	180	0.080	2.44	8.46
1st Quartile	3020	4.280	25.30	19.82
Median	4060	5.380	33.45	28.40
Mean	3700	5.058	31.36	26.23
3rd Quartile	4530	6.200	38.70	32.78
Maximum	5940	7.700	49.54	39.38

In figure 1 sudden changes in trends of solar plant generation with respect to time is shown. From the graph of data is not visualised stationary. Further we check the stationarity of data individually for both models with appropriate using of R. Next we use unit root test to check the stationarity taking first difference of time series. Table 2 shows the minimum and maximum values, mean, median and quartile of data.

5.1 Construction of ARIMAX model

The ARIMAX model follows the Box –Jenkins methodology of three stage modelling approach which is identification, estimation and diagnostic checking. In figure 1, it is shown that our data has trend and fluctuation . Before applying the ARIMAX model it is necessary that data should be stationary.For checking staionarity we apply unit root test. Next we plot the ACF(Autocorrelation function) and PACF (Partial Autocorrelation function) for identification and estimation of lags of parameters for ARIMAX model. At last we find the best fitted model, then we find the residuals for diagnostic checking. A brief description is given as follows.

5.1.1Stationary test

The unit root test is applied to consider whether our data is stationary or non stationary. Augmented Dickey–Fuller test (ADF) is used genrally for unit root in a time series. It is an augmented version of the Dickey Fuller test for a larger and more complicated set of time series models. The Augmented Dickey–Fuller (ADF) statistic, is founded as a negative. The more the number is negative, the stronger possibility of the rejection of the hypothesis that there is a unit root at some level of confidence. To avoid the fluctuation of data, the unit root test is applied to the solar power generation with different variables and results of ADF test are obtained as given in following table :

Table 3: Results of ADF test

Variable	Value of ADF test at
Solar Power	-1.8583
Solar Radiation	-1.6884
Module Temprature	-0.4504
Air Temprature	-0.5979

5.1.2 Parameter Estimation

Parameter estimation and effect of seasonality is done by studying the graphical presentation of Autocorreation Function (ACF) and Partial Autocorrelation Function (PACF) of data. Autocorrelation plots are useful tool in determining the stationarity of the given series. These plots are also help for in selecting the order of parameters for any model. If the series is correlated with its lags then, generally, there are some trend or seasonal components present in data set and therefore its statistical properties are not constant over time.ACF plots display correlation between a series and its lags. ACF plots helps in determining the order of the MA *(q)* model whereas partial autocorrelation plots (PACF) determines the order of the *AR(p)* model. The plotting of data in R gives plots 95% significance boundaries which shown by the dotted line of corresponding figures. In our time series data, we observed data there are significant autocorrelations with many lags which is shown by the ACF and PACF plots in figure 2.

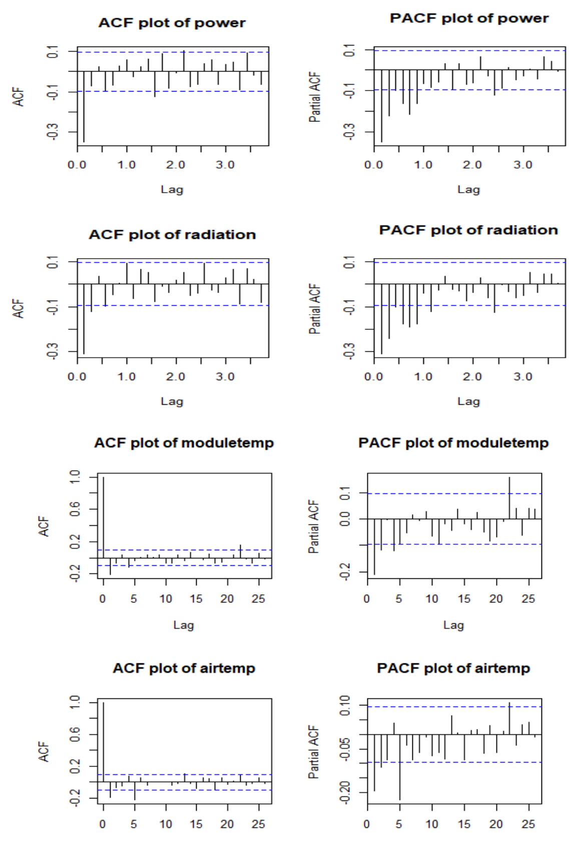

Figure:2 ACF and PACF plot of different variables

5.1.3 Model Selection Criteria

In this section we provide the best fitted model on the basis of power generation and its effect on radiation, moduletemp and airtemp. To start with the ARIMAX model, variable selection procedure consists of two steps. The first step involves identification of optimal AR and MA orders i.e p and q. In second step, we use the obtained model from the automated search and the optimal leading indicators are selected through a stepwise backward elimination procedure on the basis of of minimizing AIC.

Table 4: Estimated Value of ARIMAX (5,1,0)

ARIMAX(5,1,0)	AR1	AR2	AR3	AR4	AR5	Radiation	Moduletemp	Airtemp
Autoregressive(AR) Value	-0.7249	-0.539	-0.388	-0.217	-0.073	733.056	2.9688	-8.8005
Standard Errors	0.0493	0.0608	0.0640	0.0622	0.0519	21.0172	7.3723	9.3594
	log likelihood = -3118.57						AIC= 6253.13	

5.1.4 Residual Diagnostic Checking

Diagnostic checking is a standard tool for identification of models before forecasting the data. After the model selection procedure we will go for assumption process of the white noise of data on the basis of its residual. We know that the residuals from any selected model are assumed to be independent and usually normally distributed. Following figure 3 shows the plot of residuals on ACF and PACF the spikes in the figures shows the level of significance. From the graph it is clear that there is no external outliars available and are free from the effects of forecast values.

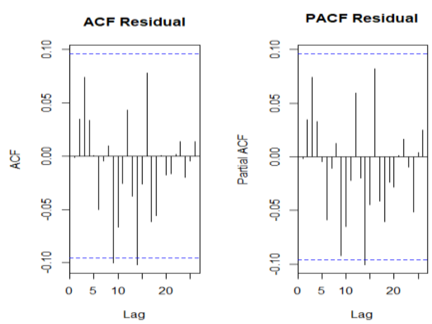

Figure 3: ACF and PACF residual of ARIMAX model

5.2 Construction of VAR Model

The vector autoregressive (VAR) model is one of the most successful, flexible, and easy to use models for the analysis of multivariate time series. It is natural extension of the univariate autoregressive model to dynamic multivariate time series. In this model we deal with the estimation procedures followed sequentially in order to estimate the cointegration VAR models to our data. The construction of VAR model is established as follows.

5.2.1 Augument Dicky fuller (ADF) test

In this section we study the stationarity, non stationarity of data and we outlines a procedures for the estimation of unit roots of the variables. By applying the Augument Dicky Fuller (ADF) test we obtained the critical values and test values for various variables and their first difference is taken for our study sake in this paper. Table 5 shows the Augument Dicky Fuller test results.

Table5 : Unit root test resuls

Variable	Lags	Deterministic terms	Critical Values			Test Values
			1%	5%	10%	
Power	2	Constant,trend	-3.98	-3.42	-3.13	-6.5315
Differencepower	1	Constant	-3.44	-2.87	-2.57	-20.5235
Radiation	2	Constant,trend	-3.98	-3.42	-3.13	-6.4043
Differenceradiation	1	Constant	-3.44	-2.87	-2.57	-21.1541
Moduletemp	2	Constant,trend	-3.98	-3.42	-3.13	-2.5493
Differencemoduletemp	1	Constant	-3.44	-2.87	-2.57	-18.1512
Airtemp	2	Constant,trend	-3.98	-3.42	-3.13	-2.5277
Differenceairtemp	1	Constant	-3.44	-2.87	-2.57	-17.5643

From the above table it is clear that ADF test statstics values are higher than the critical values at 0.01, 0.05 and 0.10 siginificance level.

5.2.2Lag Optimisation and Co-integration test

In this section the lag optimisation and co-integration is obtained by using the Philips Outliars method for outlines the VAR model. The information criteria (IC) method is used for right lag order selection by minimising the mean square error. We estimate our VAR model for solar generation series and for this the order of p must specified. The estimation of parameters is obtained by commonly used criteria Akaike(AIC), Schwarz (BIC), Hannan-Quinn(HQ) and at last the Final perdiction error (FPE) which is shown by following table:-

Table6: Estimation of VAR model

Criteria	Order of parameter
AIC(n)	5
HQ(n)	2
SC(n)	1
FPE(n)	5

From the above table it is clear that the optimal lag number is p = 5 in case of AIC and FPE, whereas the HQ criterion indicates p = 2 and the SC criterion indicates p = 1 as optimal lag length.

Phillips Ouliaris co-integration test are based on residuals and the tests are the variance ratio test and the multivariate trace statisitics. These residual-based tests are not much different from the unit root tests. The only difference is the data are residual from the cointegration [17]. These tests checked the null hypothesis of no cointegration against the alternative hypothesis of the presence of cointegration. In our analysis of solar power of generation the value of co-integration test provides eviedence that the series are cointegration because the null hypothesis is rejected at the 1% level.

5.2.3Estimation of VAR model and Diagonastic checking

The model order determination is an important step in vector autoregressive (VAR). If the model of the selected order is most accurate among the estimated models then we use can say that the performance of the order selection is optimal. Waele and Broerson in their paper [21] elaborated an order selection criteria for VAR model. The following tables showed the estimated results of VAR(5) model of power, radiation, module temperature and air temprature as according to our solar power data.

Table 7:Estimation results for equation Power

Parameter	Estimate	S.E.	t value	Pr> ?	Sig.
power.11	-0.5917	0.1065	-5.553	5.19e-08	***
radiation.11	106.3364	91.0188	1.168	0.24340	
moduletemp.11	17.7448	18.0279	0.984	0.32558	
airtemp.11	-46.5298	23.8478	-1.951	0.05176	.
power.12	-0.3915	0.1282	-3.054	0.00242	**
radiation.12	92.5381	107.9505	0.857	0.39184	
moduletemp.12	-8.6383	18.5481	-0.466	0.64167	
airtemp.12	-19.0008	24.1006	-0.788	0.43094	
power.13	-0.4305	0.1330	-3.237	0.00131	**
radiation.13	215.2670	114.3117	1.883	0.06042	.
moduletemp.13	-7.2956	18.6667	-0.391	0.69613	
airtemp.13	-18.3185	24.1746	-0.758	0.44905	
power.14	-0.2694	0.1236	-2.179	0.02991	*
radiation.14	37.3870	105.4106	0.355	0.72302	
moduletemp.14	-13.0664	18.5609	-0.704	0.48187	
airtemp.14	-5.0830	24.1172	-0.211	0.83318	
power.15	-0.15.04	0.1018	-1.477	0.14045	
radiation.15	53.0399	87.1267	0.609	0.54303	
moduletemp.15	-56.6639	18.0634	-3.137	0.00184	**
airtemp.15	-15.9026	23.6518	-0.672	0.50175	

Table 8:Estimation results for equation Radiation

Parameter	Estimate	S.E.	t value	Pr> ?	Sig.
power.11	0.0001502	0.0001288	1.166	0.2442	
radiation.11	-0.5851435	0.1099913	-5.320	1.75e-07	***
moduletemp.11	0.0235253	0.0217857	1.080	0.2809	
airtemp.11	-0.0393719	0.0288187	-1.366	0.1727	
power.12	0.0001396	0.0001549	0.901	0.3682	
radiation.12	-0.4457244	0.1304523	-3.417	0.0007	***
moduletemp.12	0.0105276	0.0224144	0.470	0.6388	
airtemp.12	-0.0242287	0.0291243	-0.832	0.4060	
power.13	-0.0001712	0.0001607	-1.066	0.2873	
radiation.13	-0.0451261	0.1381394	-0.327	0.7441	
moduletemp.13	-0.0030539	0.0225577	-0.135	0.8924	
airtemp.13	-0.0377532	0.0292137	-1.292	0.1970	
power.14	-0.0001181	0.0001494	-0.791	0.4296	
radiation.14	-0.1345920	0.1273830	-1.057	0.2913	
moduletemp.14	-0.0188866	0.0224298	-0.842	0.4003	
airtemp.14	-0.0031244	0.0291443	-0.107	0.9147	
power.15	-0.0001331	0.0001230	-1.082	0.2799	
radiation.15	0.0084801	0.1052879	0.081	0.9358	
moduletemp.15	-0.0546320	0.0218287	-2.503	0.0127	*
airtemp.15	-0.0230914	0.0285820	-0.808	0.4196	

Table9:Estimation results for equation Moduletemp

Parameter	Estimate	S.E.	t value	Pr> ?	Sig.
power.11	0.0002558	0.0003515	0.728	0.467285	
radiation.11	-0.2952637	0.3002905	-0.983	0.326087	
moduletemp.11	-0.2149203	0.0594778	-3.613	0.000342	***
airtemp.11	-0.0445392	0.0786789	-0.566	0.571658	
power.12	0.0002625	0.0004230	0.620	0.535298	
radiation.12	-0.2930487	0.3561517	-0.823	0.411112	
moduletemp.12	-0.0865873	0.0611941	-1.415	0.157876	
airtemp.12	-0.1303315	0.0795131	-1.639	0.101993	
power.13	-0.0003885	0.0004388	-0.885	0.376441	
radiation.13	0.3206212	0.3771385	0.850	0.395767	
moduletemp.13	0.0098178	0.0615855	0.159	0.873423	
airtemp.13	-0.1360134	0.0797572	--1.705	0.088925	.
power.14	-0.0001578	0.0004079	-0.387	0.699069	
radiation.14	0.0449452	0.3477720	0.129	0.897236	
moduletemp.14	-0.0924651	0.0612364	-1.510	0.131859	
airtemp.14	0.0597975	0.0795676	0.752	0.452786	
power.15	-0.0004042	0.0003358	-1.204	0.229493	
radiation.15	0.3723336	0.2874496	1.295	0.195981	
moduletemp.15	-0.1135245	0.0595951	-1.905	0.057523	.
airtemp.15	-0.0189182	0.0780325	-0.242	0.808566	

Table10:Estimation results for equation Air temperature

Parameter	Estimate	S.E.	t value	Pr> ?	Sig.
power.11	9.391e-05	2.539e-04	0.370	0.711649	
radiation.11	-3.994e-02	2.169e-01	-0.184	0.853979	
moduletemp.11	1.922e-02	4.295e-02	0.447	0.654860	
airtemp.11	-2.145e-01	5.682e-02	-3.775	0.000185	***
power.12	7.565e-05	3.055e-04	0.248	0.804535	
radiation.12	-2.726e-02	2.572e-01	-0.106	0.915656	
moduletemp.12	-2.539e-02	4.419e-02	-0.575	0.565904	
airtemp.12	-1.324e-01	5.742e-02	-2.305	0.021688	*
power.13	-5.109e-05	3.169e-04	-0.161	0.871989	
radiation.13	1.732e-01	2.724e-01	0.636	0.525075	
moduletemp.13	-4.337e-02	4.448e-02	-0.975	0.330104	
airtemp.13	-1.063e-01	5.760e-02	-1.845	0.065753	.
power.14	2.539e-04	2.945e-04	0.862	0.389292	
radiation.14	-9.341e-02	2.511e-01	-0.372	0.710152	
moduletemp.14	-2.172e-02	4.442e-02	-0.491	0.623599	
airtemp.14	-2.587e-03	5.746e-02	-0.045	0.964114	
power.15	-3.804e-04	2.425e-04	-1.568	0.117616	
radiation.15	3.330e-01	2.076e-01	1.604	0.109482	
moduletemp.15	3.137e-02	4.304e-02	0.729	0.466553	
airtemp.15	-2.575e-01	5.635e-02	-4.570	6.55e-06	***

Level of significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Next we daignosis the white noise from VAR model. The following figure 4 shows the plot of correlogram of residuals models.

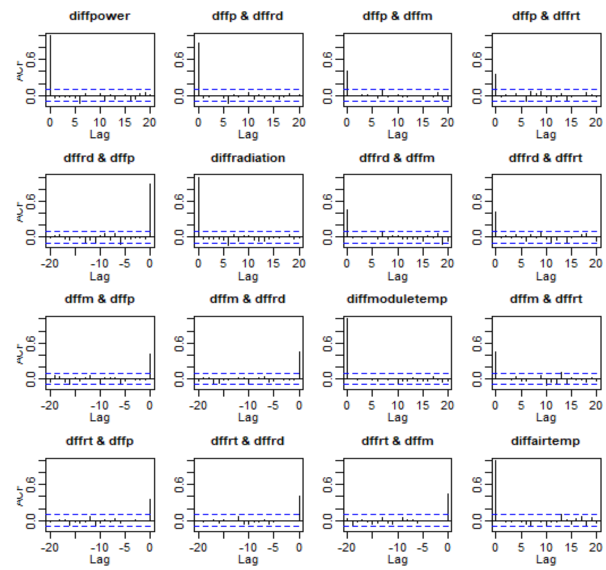

Figure 4: ACF Plots of residuals model VAR(5)

The plots along the diagonal are the individual ACF for each model's residual. In addition , The other plots are the cross-correlation plots of each set of residuals and these would also resemble white noise, however cross-correlation.

5.2.4 Granger Casulity test

VAR models describes the joint generation process of a number of variables, which are used for investigating relations between the variables. A special type of relation was found by Granger (1969), and is known as Granger casulity test. This test is used to determine the relationship between one variable with other variables that effect each other. The relationship between the above variables are illustrated in our following table:

Table 11 : Relationship between variables

Dependent	**p-Value of H_0 Power**	**p-Value of H_0 Radaition**	**p-Value of H_0 Moduletemp**	**p-Value of H_0 Airtemp**
Power	-	0.4478(Accept)	0.0175(Reject)	0.3261(Accept)
Radiation	0.267(Accept)	-	0.1574(Accept)	0.3638(Accept)
Moduletemp	0.977(Accept)	0.9048(Accept)	-	0.2707(Accept)
Airtemp	0.797(Accept)	0.9583(Accept)	0.8616(Accept)	-

In Granger casulity test the H_0 (Null hypothesis) is taken as that there exist no Granger cause when probability value is more then 5% level of significance i.e null hypothesis is accepected, otherwise rejeceted. In this study it is clear from the above table that there is only one Granger cause relationship exists between power and module temperature as it is rejected at that level.

5.1.3 Comparision and Forecast

Based on our experimental data studied under ARIMAX and VAR model, the results can be used to forecast the solar plant generation. The solution of the best method can be done by the method of comparison based on the error measures. In the following table three error measures are calculated namely Mean Absolute Error (MAE), Mean Absolute Scaled Error (MASE) and Root Mean Square Error (RMSE) for forecasting the best accurate model under study on the basis of the minimum error values.

Table12: Error Measures of two models

Error Measures	MAE	MASE	RMSE
ARIMAX	241.8104	0.8919	451.118
VAR	**0.8285**	**0.70368**	**1.12757**

From the above table it is clear that VAR model is the best model as it acquires minimum values for all the three measures. Now in the following table we forecasted the lowest and highest value of our solar plant generation up to 200 days at the 95% significance level.

Table13 : Forecast value of Solar plant generations different variables

Signifiacance level→ **Variable↓**	**Low95%**	**High95%**
Power	**2132.5986**	**5880.207**
Radiation	**3.308868**	**7.840227**
Module temperature	**32.73838**	**44.87894**
Air temperature	**16.95535**	**25.89984**

Again in the following figure it is clear that forecasting of our solar plant generation data is provided with each variable.

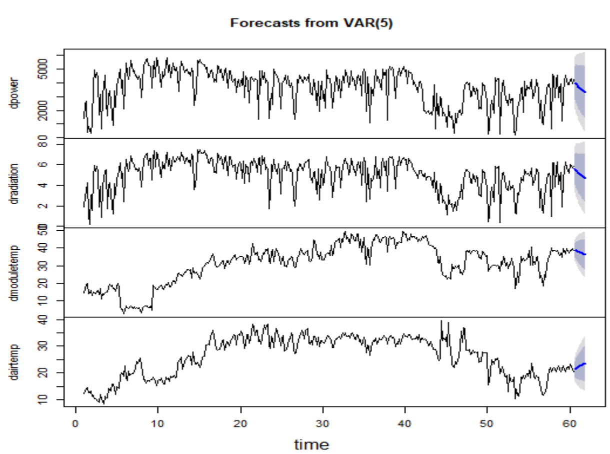

Figure 5: Forecast of Solar plant generation with effect of different variables

From the above figure we can see that the forecast trend is decreasing for power and radiation. In case of module temperature the trend is moderate but for air temperature the trend is seems to be increasing. By studying this trend we can come to the decision that seasonal and weather condition effects our forecast trend as Punjab is a state of India where it experiences the extreme weather in each season for example too hot in summers (approx. 39.38 C^{0}) and too cold in winters (approx. 8.46 C^{0}). So these weather condition effects on our forecast trend on power. Again in this study it is clear from Table 11, that module temperature effects on power and our forecast trend is also shows the same effect. In addition, we can also come to the conclusion that not only the seasional and weather conditions effets our trend of data but factors like technical fault also effect on production trend of power generation. So in our present study we take the VAR model where highest and lowest forecast values up to 95% level of significance is considered for better forecast of our existing data set.

VI. CONCULSION AND FUTURE SCOPE

The use of solar energy is becoming popular and necessary in our day to day life. Country like India is still unbalanced in electricity production. Day by day our consumption of electricity is increasing but the production is not comparatively increasing. There are various sources for production of electricity like coal, hydro, wind, solar, nuclear etc. But the solar power has abundant source of energy and India has suitable climate for producing ample solar energy. The power generation produced from hydro plant is not regular due to irregular flow of water and production of electricity from coal is not sufficient due to limited sourse of fossil fuels. That is the reason recently india is opting producing solar power energy and getting positive response as the power consumption is increasing frequently and efforts are given for future production prospects. This is the reason our research is focused on forecasting of solar plant generation which will be helpful to forecast the solar plant generation data and also this information can be used in various Government and private sectors for producing solar power energy. In this paper we focused on forecasting performance of ARIMAX and VAR models for solar plant generation data. Based on the results of error measures , it is concluded that VAR model is better for forecasting the different variables of solar plant generation as compared with ARIMAX method. In our data analysis, we also forecast the minimum and maximum forecast value by using VAR(5) model at 95% level of significance which can be used by various power sectors to distribute power, which is shown in table 7. In this paper we studied the solar forecasting methods and evaluation metrics is also discussed.

So at last we can say that as accurte we forecast the weather, the more accurate we can forecast the future prediction of solar plant generation data. As we have taken solar plant generation of data of a power plant of Phullokari village,Punjab and Punjab has an extreme weather condtion like too hot in summer and too cold in winters, so power generation of data also experiences a high fluctuation of solar generation data. So this is the limitation faced by this region for stable generation of solar power. Government launches various schemes for production of solar power but we need to take some initiative for using of renewable energy as much as at a place of conventional energy sources. Moreover solar power plant generation is also can be taken as alternative source of power generation as compared to thermal power plants in Punjab.

REFERENCES

[1] A. Jalalkamali, M. Moradi and N. Moradi, "Application of several artificial intelligence models and ARIMAX model for forecasting drought using Standardized Precipitation Index", Int. J. Environ. Sci. Technology, Vol. 12, pp. 1201-1210, 2015.
[2] Billy M. Williams, "Multivariate vehicular traffic flow prediction, Evaluation of ARIMAX modelling", Journal of the transportation research board ,pp. 194-200, January 2001
[3] B. Yogarajah, C. elankumaran and R. Vigneswaran, "Application of ARIMAX Model for forecasting Paddy Production in Trincomalee District in Sri Lanka", Proceeding of third international Symposium, Vol. 24, pp. 2-3, July 2013
[4] Christopher Bennett, Rodney A. Stewart and Junwei Lu, " Autoregressive with Exogenous Variables and Neural Network Short- Term Load Forecast Models for Residential Low Voltage Distribution Networks", Energies, Vol. 7, pp. 2938-2960, 2014
[5] Ding N, Besanger Y, " Time series method for short – term load forecastingusing smart metering in distribution systems, " International Proceeding of the IEEE Trondheim Power Tech, pp. 1-6, 2011
[6] D.K.Chaturvedi and Isha, " Solar Power Forecasting: A Review",International Journal of Computer Applications", Vol.145, Issue. 6, July 2016
[7] El Mostafa Bentour, " A ranking of VAR and structural models in forecasting",http://mpra.ub.uni-muenchen.de/61502/, January 2015
[8] Herui Cui and Xu Peng , "Short –Term City Electric Load Forecasting with Considering Temprature Effects: An Improved ARIMAX Model", Hindawi Pubishing Corporation, Mathematical Problem in Engineering , Volume 2015
[9] Gulam Ali "Cointegration VAR and VECM and ARIMAX Econometric Approaches for water quality Variaties" Journal of Statistical and Econometric Methods, Vol. 4, Issue. 1, pp.1-38, 2015
[10] Harendra Kumar Yadav, Vinay Kumar Yadav and Vijay Kumar, " Potential of solar energy in India: A Review" International Advanced Research Journal of Science, Engineering and Technology, Vol. 2, Issue 1, May 2015
[11] Jabar H. Yousif and Hussein A. Kazem, " Modeling of Daily Solar Energy System Prediction using Soft Cmputing Methods for Oman", Research Journal of Applied Sciences, Engineering and Technology 13(3): 237-244, 2016
[12] Jai Singh Arya, Aadesh Kumar Arya and Sanjeev Aggarwal, " Recent Trends in Solar Energy Generation", Association of Computer Electronics and Electrical Engineers, 2013
[13] Javed Iqbal, "Forecasting Methods: A Comparative Analysis", Proc. Eight Stat. Sem. 2001, K.U, MPRA Paper No.23856, posted29. Auguest 2013
[14] Jing Fan, Rui Shan and Xiaoqin Cao & Peiliang Li, "The Analysis to Tertiary- industry with ARIMAX Model", Journal of Mathematics Research, Vol.1, No.2, September 2009
[15] Mohamed Abuella and Badrul Chowdhary, "Solar power Forecasting Using Support Vector Regression" in proceeding of the American Society for Engineering Management 2016
[16] Paul S. P. Cowperwait and Andrew V. Metcalfe,"Introductory Time Series with R" ,Springer, 2008
[17] Phillips, P. C. B.; Ouliaris, S. (1990). "Asymptotic Properties of Residual Based Tests for Cointegration".Econometrica. Vol.58, Issue.1, pp. 165–193.
[18] Renny Elfira Wulansari, Setiawan and Suhartono, "A Comparison of Forecasting Performance of Seasonal ARIMAX and Hybrid Seasonal ARIMAX –ANN of Surabaya's Currency Circulation Data", International Journal of Management and Applied Science , Vol.2, Issue-10,Oct-2016
[19] Robert H. Shumway and David S. Stoffer, "Time Series Analysis and Its Applications With R Examples",Second Edition, Springer text in Statistics, 2005
[20] S.M. Husnain Bokhari and Mete Feridun, "Forecasting inflation through Econometric models : An Empirical Study on Pakistani Data", Vol. 7, Issue. 1, pp.39-47, 2006
[21] Stijin de Waele and Piet M.T. Broerson, " Order Selection for Vector Autoregressive Model" ,IEEE transactions on signal processing, Vol. 51, Issue. 2, February, 2003
[22] Uyodhu Amekauma Victor- Edema and Prof Isaac Didi Essi , "Autoregressive Integrated Moving Average with Exogenous Variable(ARIMAX) Model for Nigerian Non Oil Export", European Journal of Business and Management, Vol. 8, Issue. 36, 2016
[23] Wiwik Anggraeni, Kuntoro Boga Andri, Sumaryanto and Faizal Mahananto, "The performance of ARIMAX Model and Vector Autoregressive (VAR) Modelforecasting strategic commodity price in Indonesia", Procedia Computer Science, pp.189-196, November 2017
[24] Wiwik Anggraeni, Retno Aulia Vinarti and Yuni Dwi Kurniawati, "Performance Comparisions Between Arima and arimax Method in Moslem Kids Clothes Demand Forecasting: Case Study", Procedia Computer Science, Vol.72, pp.630-637, 2015

Authors Profile

Dr. Madhuchanda Rakshit, Ph.D in Operation research. Presently working as Assistant Professor at Guru Kashi University. She has more than 25 research paperson her own credit. Her area of research is Queneing theory,Time series analysis and Quadratic and fractional programming problems.

Sukhpal Kaur, M.Sc. Mathematics from Punjabi University, Patiala. Presently a research scholar of Applied Science Department, Guru Kashi University. She has published three research paper in international journals and presented few research paper related to her thesis work in various National and International conferences. Her area of research is time series models

