



# Solar Energy Data Modelling



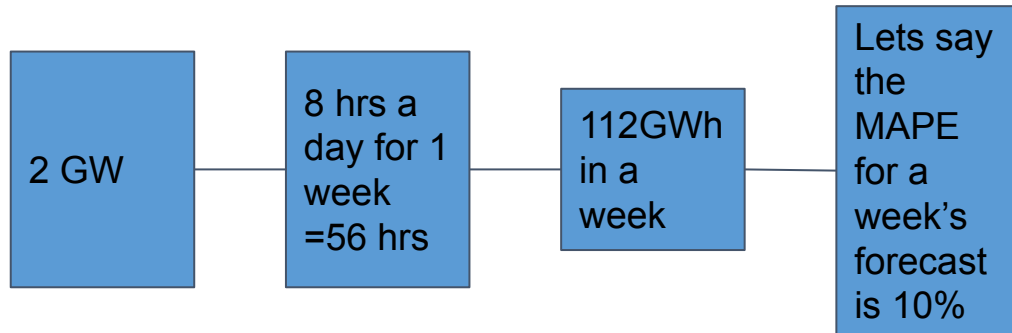
**BITS Pilani**

Pilani Campus

NAME : Raj Shree Singh  
ID : 2017B4A70808P  
Study Project (MATH F266)  
Department of Mathematics

Supervised by :  
Prof Sumanta Pasari

# Need For Forecasting



Thus

$$| \text{Forecasted value} - \text{Actual Value} | = 0.1 \times 112 \text{GWh} \\ = 11.2 \text{GWh}$$

According to Ministry of Power GoI, the per capita consumption of India during year 2019-20 was 1208 KWh. Thus 11.2 GWh can fulfill energy demands of 9 people for a year.



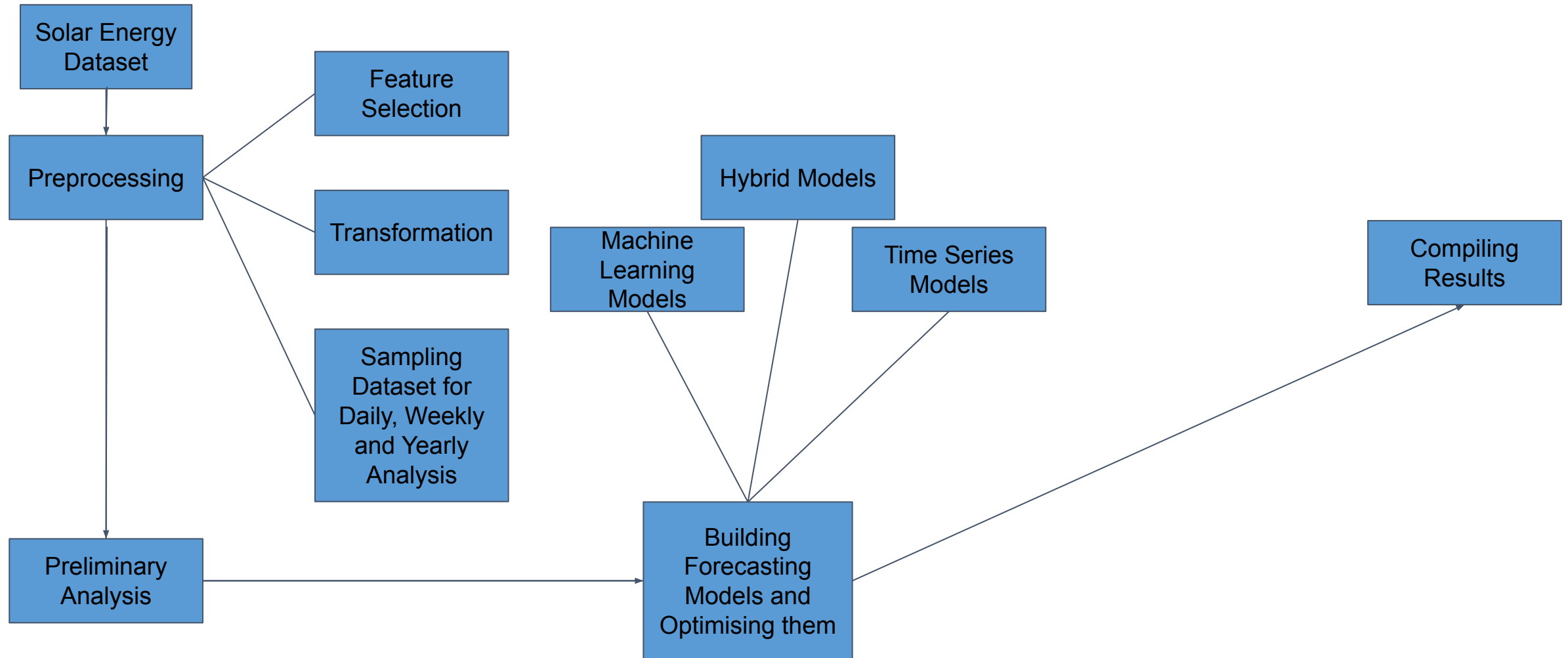
Pavagada solar park, also known as Shakti Sthala, is a 2GW solar complex developed in Pavagada, Tumkur district, approximately 180km from Bengaluru, Karnataka, India.

# AIM



**Spatio temporal analysis of Solar Energy data across  
Karnataka, Andhra, Rajasthan, Gujarat, Tamil Nadu,  
Telangana**

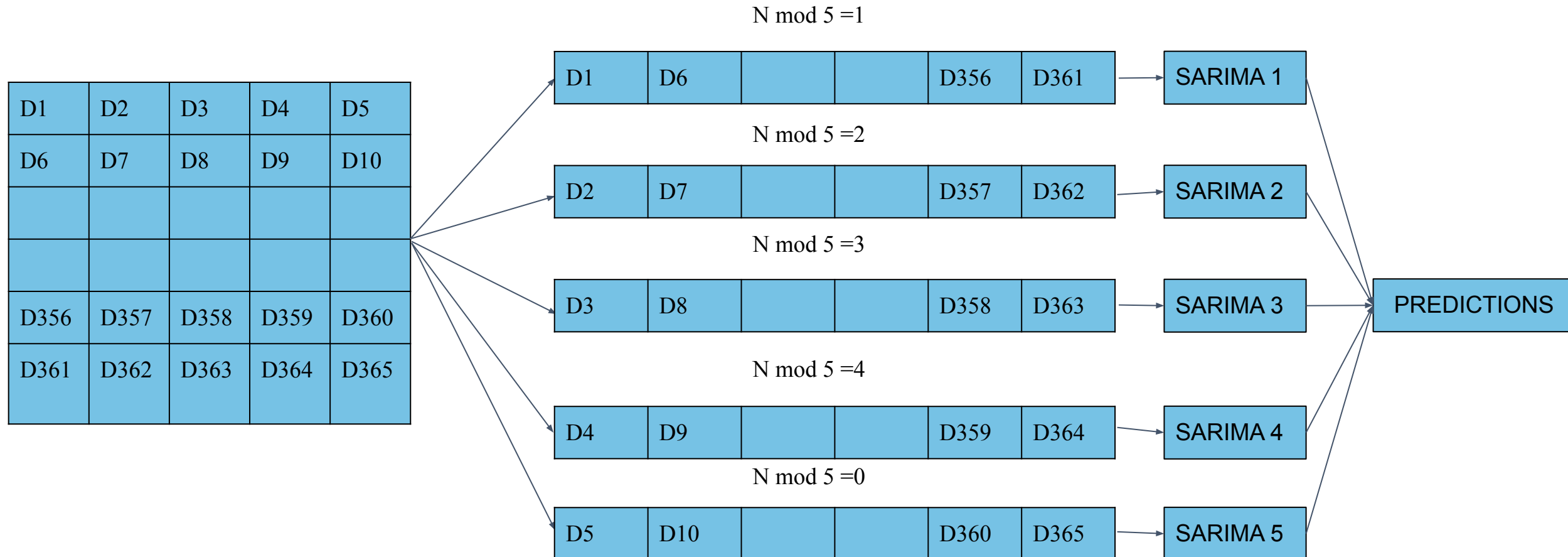
# METHODOLOGY



# Preprocessing for SARIMA



## DAILY SARIMA MODEL



# Preprocessing for ML models



The input matrix and output for the LSTM and MLP models are-

	Input Matrix	Output
<b>Monthly Dataset</b>	GHI values for the days in that month (input dim -31,28,30)	Mean GHI value for the month
<b>Weekly Dataset</b>	GHI values for the days in that week (input dim -7)	Mean GHI value for the week
<b>Daily Dataset</b>	GHI values for the 9 hours in that day (input dim -9)	Mean GHI value for that day

# Preliminary Analysis



## Dataset -

The dataset used in this project is obtained from National Solar Radiation Dataset (NSRDB) maintained by US Department of Energy. The image below show a snap of the hourly dataset of a place located near Pokhran, Rajasthan from a period of January 1 2000 to December 31 2014. This dataset is used to collect data across 6 states.

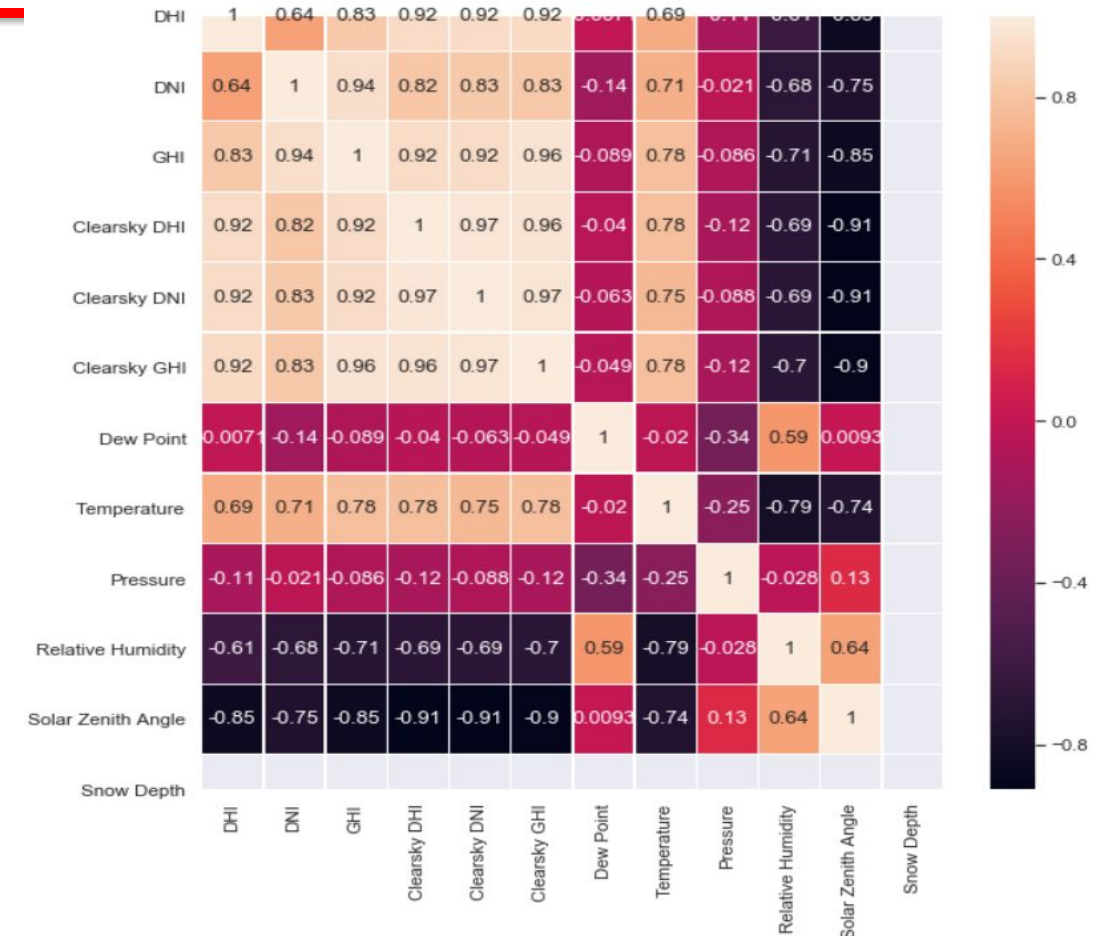
Year	Month	Day	Hour	Minute	DHI	DNI	GHI	Clearsky D	Clearsky D	Clearsky G	Dew Point	Temperatu	Pressure	Relative H	Solar Zenit	Snow Dept	Wind Spee
2000	1	1	0	0	0	0	0	0	0	0	-4	13.81144	982.7498	27.44786	174.756	0	3.976945
2000	1	1	1	0	0	0	0	0	0	0	-4	13.1173	982.6701	29.04886	169.5442	0	4.017371
2000	1	1	2	0	0	0	0	0	0	0	-4	12.43396	982.4042	31.21666	156.3409	0	3.992325
2000	1	1	3	0	0	0	0	0	0	0	-3	11.79763	982.5553	34.95607	142.9459	0	3.981812
2000	1	1	4	0	0	0	0	0	0	0	-1	11.19549	982.5581	41.17754	129.6258	0	3.985008
2000	1	1	5	0	0	0	0	0	0	0	-11	11.13673	984.3793	19.56375	116.4282	0	3.48046
2000	1	1	6	0	0	0	0	0	0	0	-10	10.64698	985.1686	21.35785	103.5838	0	3.523518
2000	1	1	7	0	0	0	0	0	0	0	-9	11.26521	985.9256	22.16892	91.18165	0	3.546341
2000	1	1	8	0	78	306	135	74	354	139	-8	14.27732	986.4865	20.12081	79.45063	0	3.323988
2000	1	1	9	0	114	597	331	121	600	339	-6	17.67728	986.8563	18.36524	68.75594	0	2.818877



# GHI and Clearsky GHI



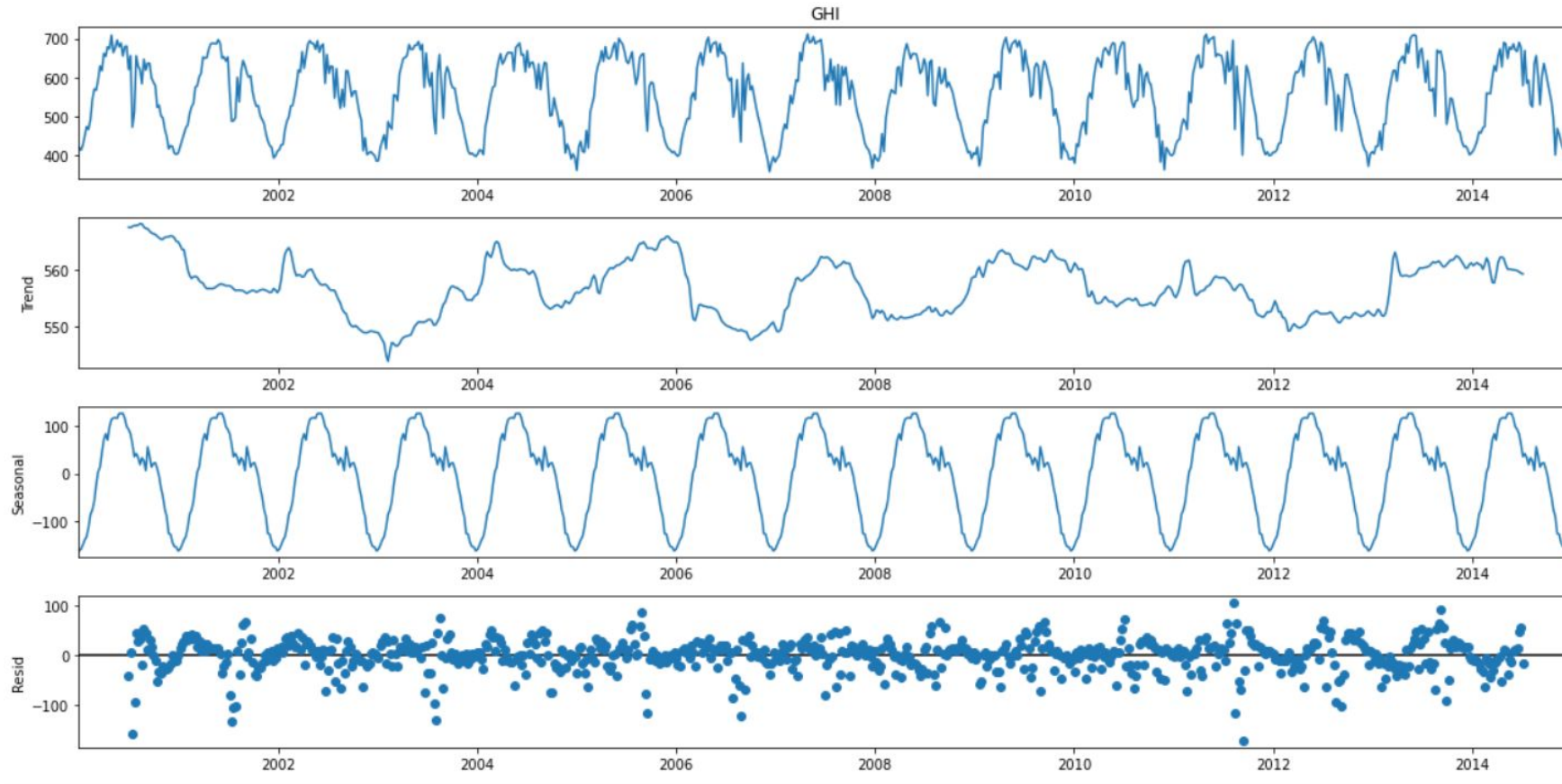
1. GHI - The total amount of short wave radiation received per unit area from above by the horizontal earth's surface is known as GHI or Global Horizontal Irradiance.
2. Clearsky GHI - The maximum solar radiation at a specific time and place on the earth's surface when no cloud is present. Reflections from clouds can result in instantaneous measurements greater than clear sky predictions.



Correlation Heatmap for a site located in Tamil Nadu

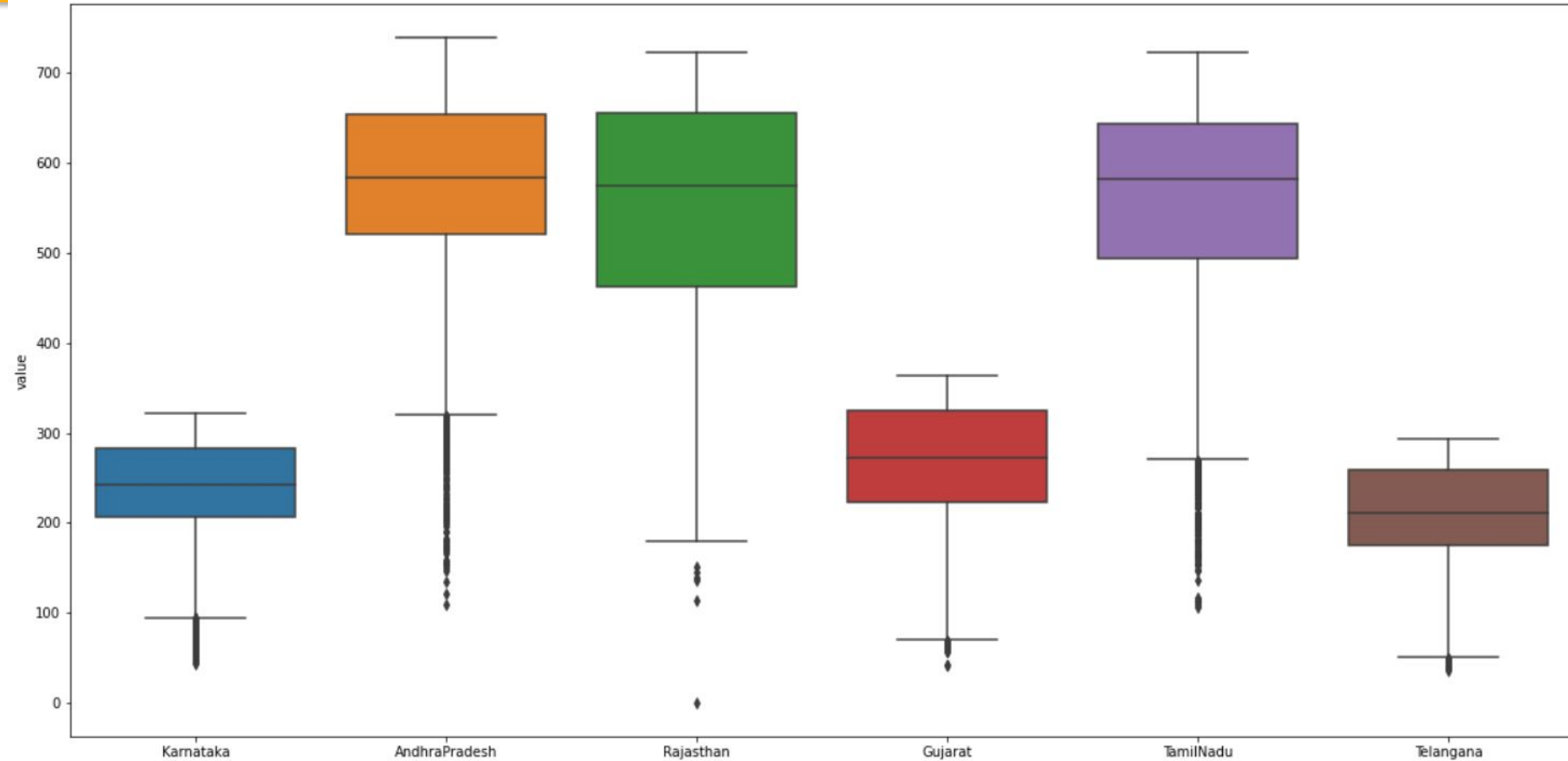


# Time Series Decomposition



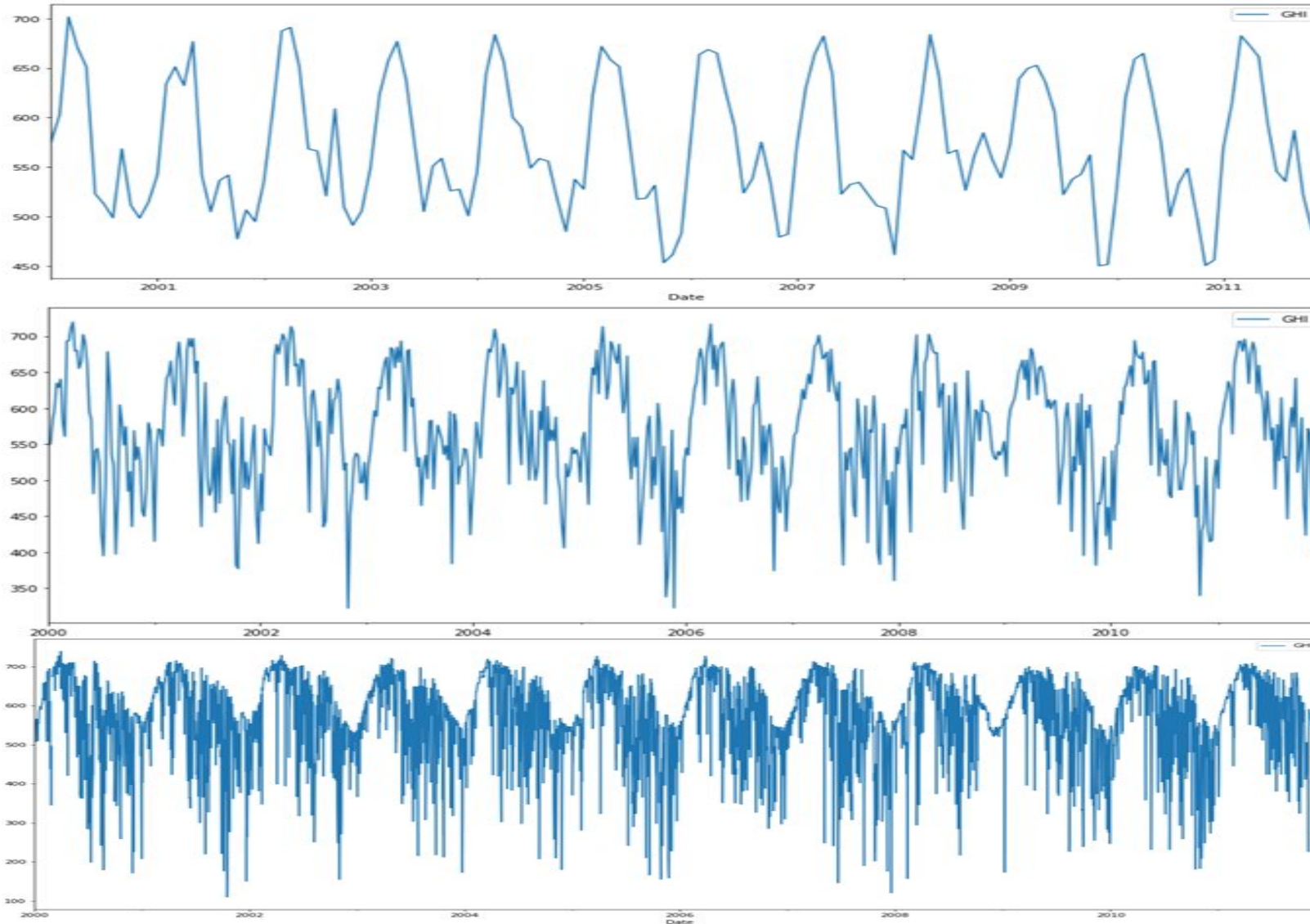
Time Series Decomposition For  
Rajasthan (Weekly)

# BOXPLOT

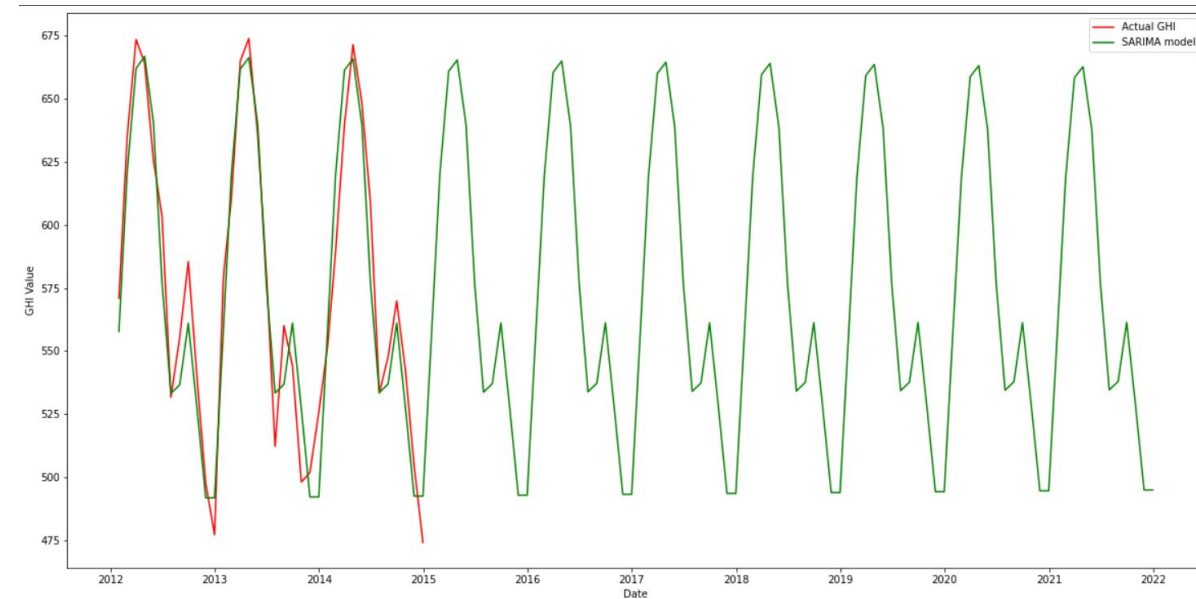


STD Deviation	54.31	106.34	110.04	61.75	117.59	54.96
Coeff. of Variance	0.23	0.186	0.197	0.229	0.212	0.262

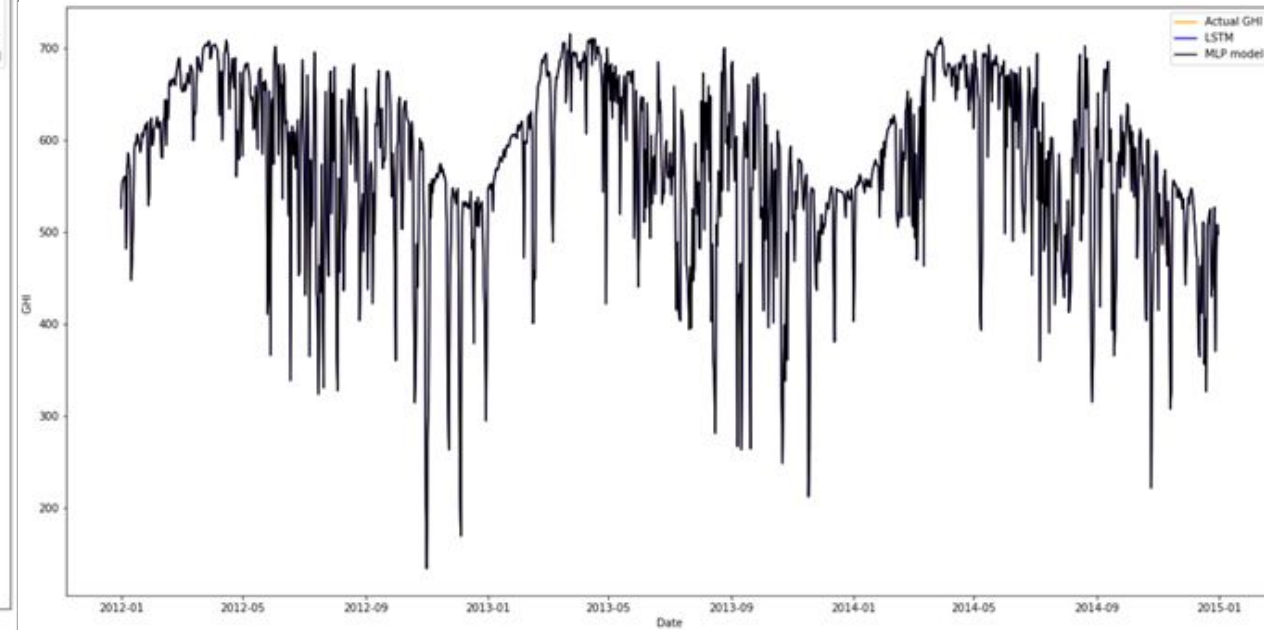
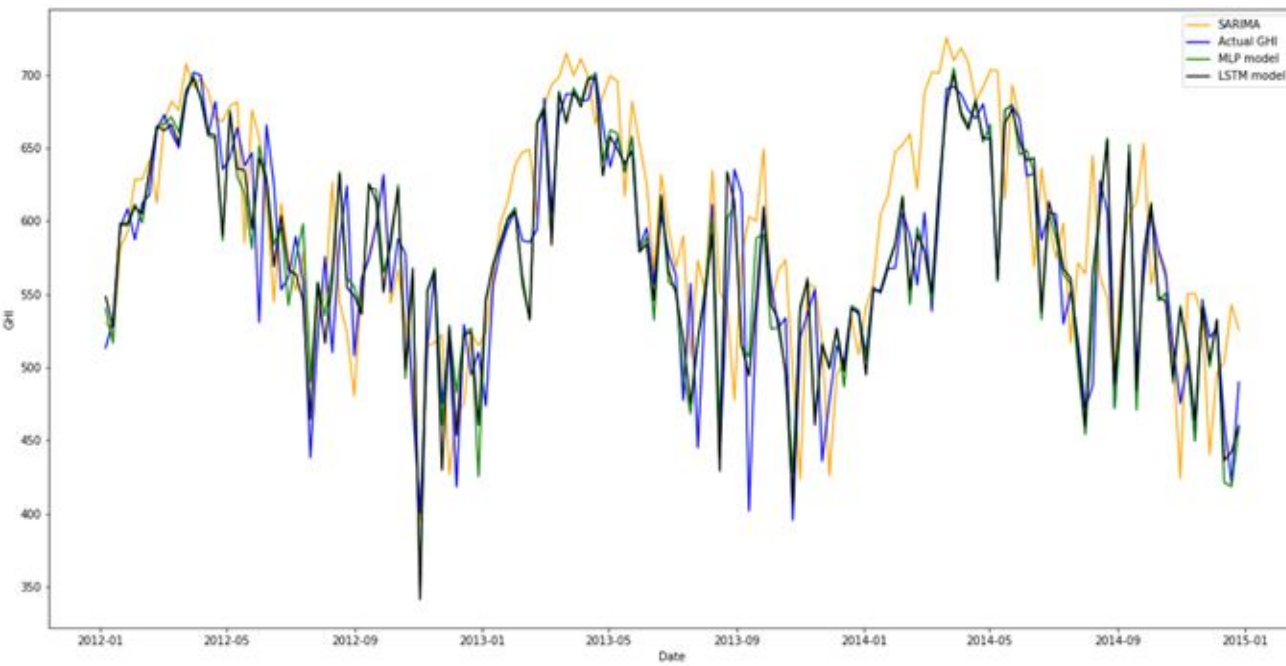
# Andhra Pradesh (Training Dataset)



# Andhra Pradesh (Monthly Predictions)

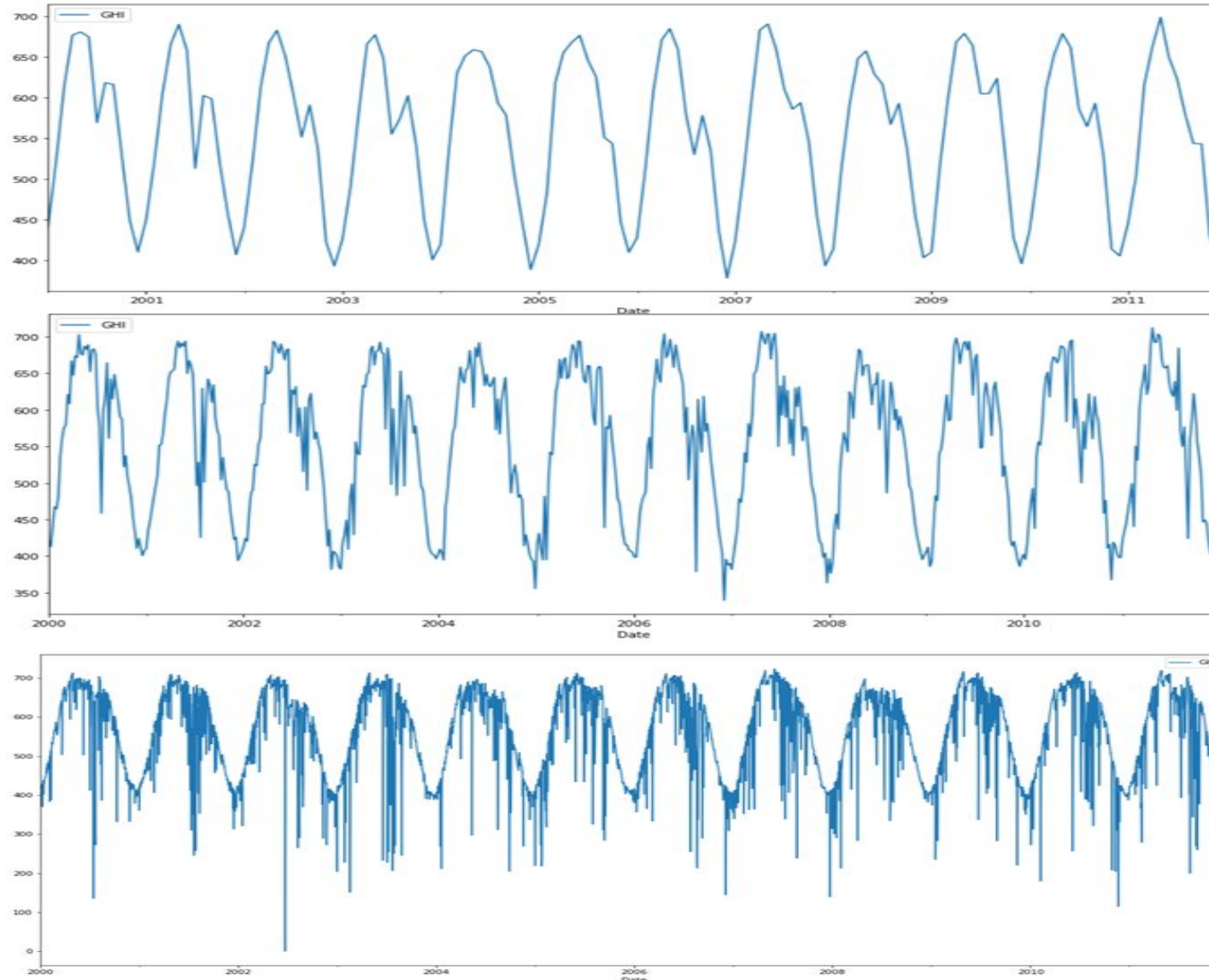


# Andhra Pradesh (Weekly and Daily)

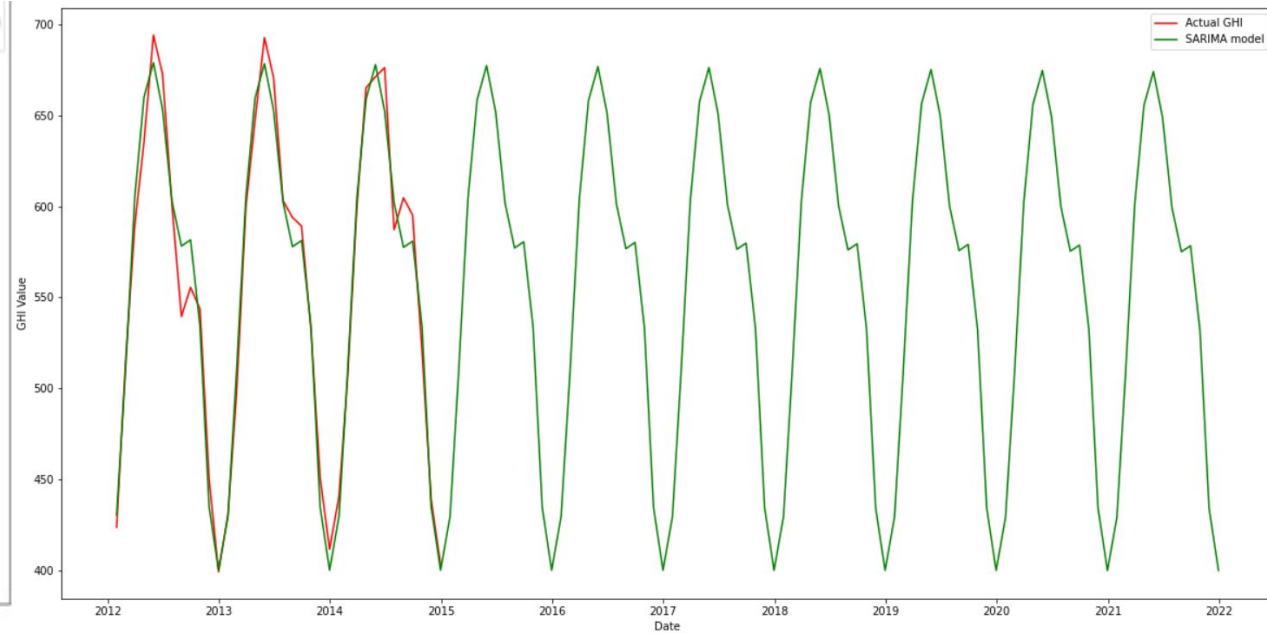
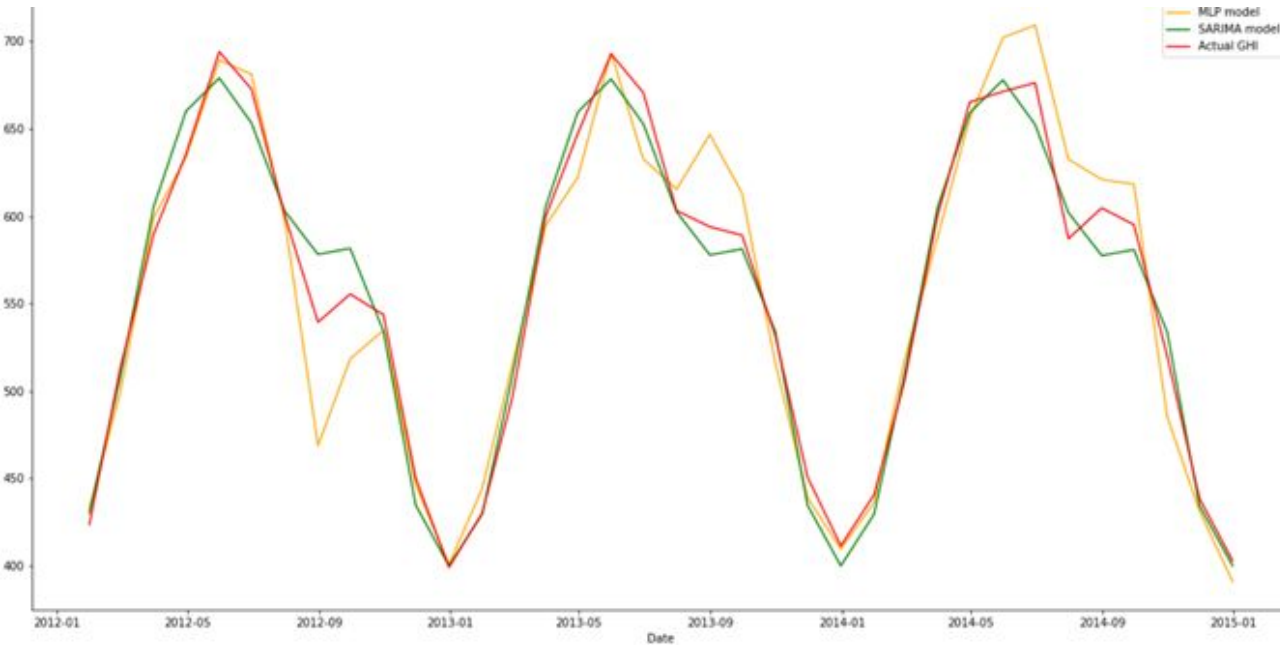




# Rajasthan (Training Dataset)

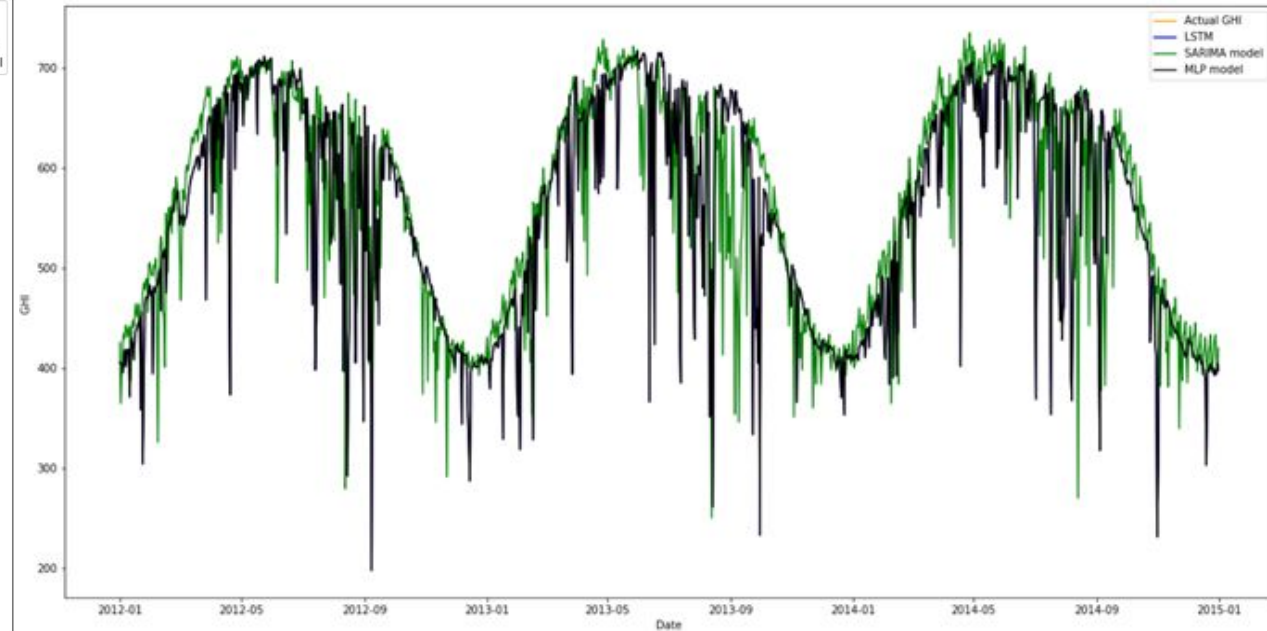
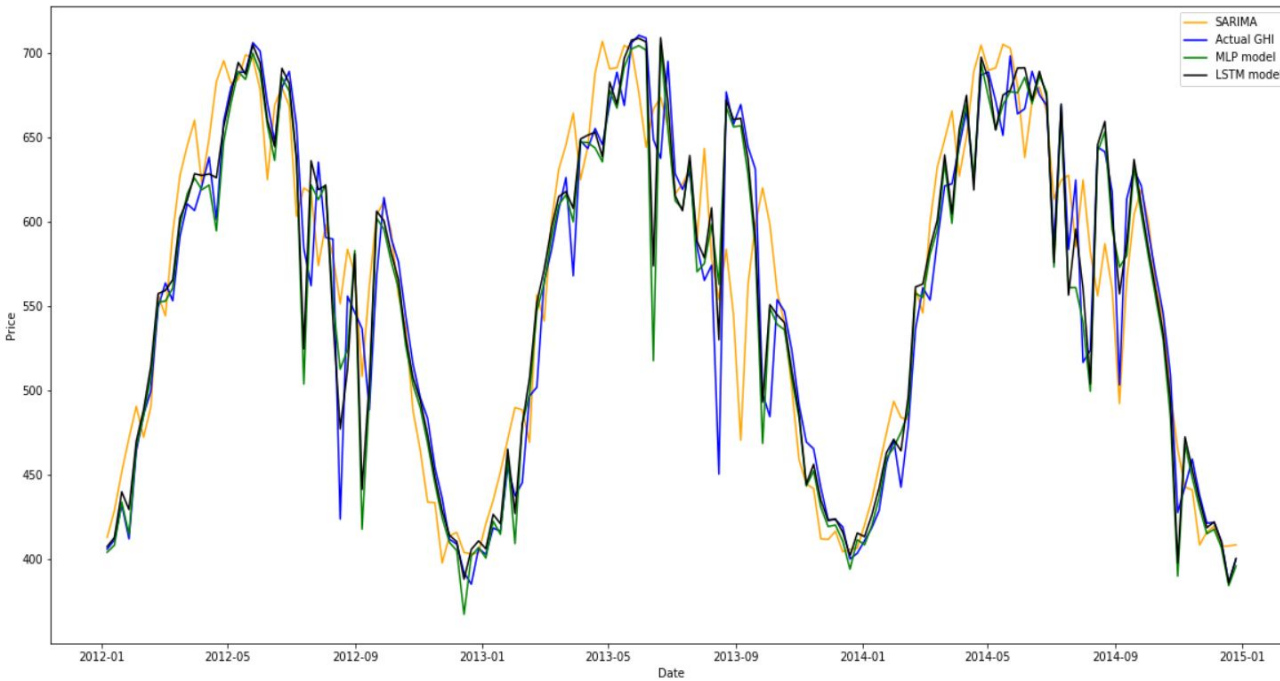


# Rajasthan (Monthly Predictions)

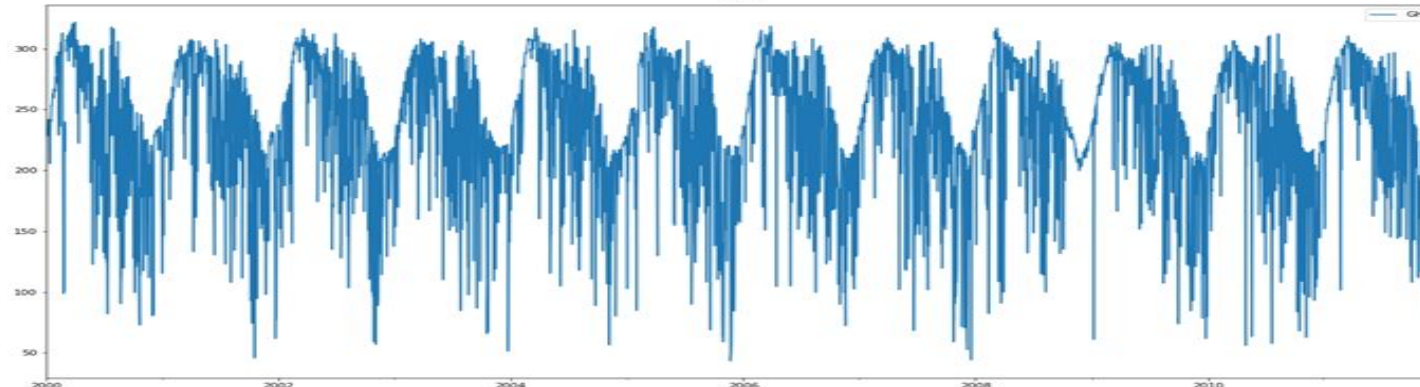
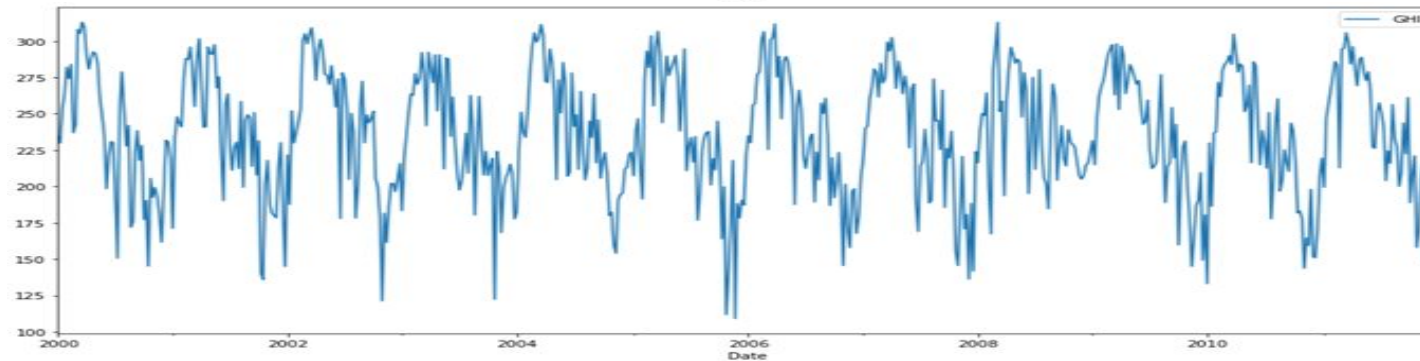
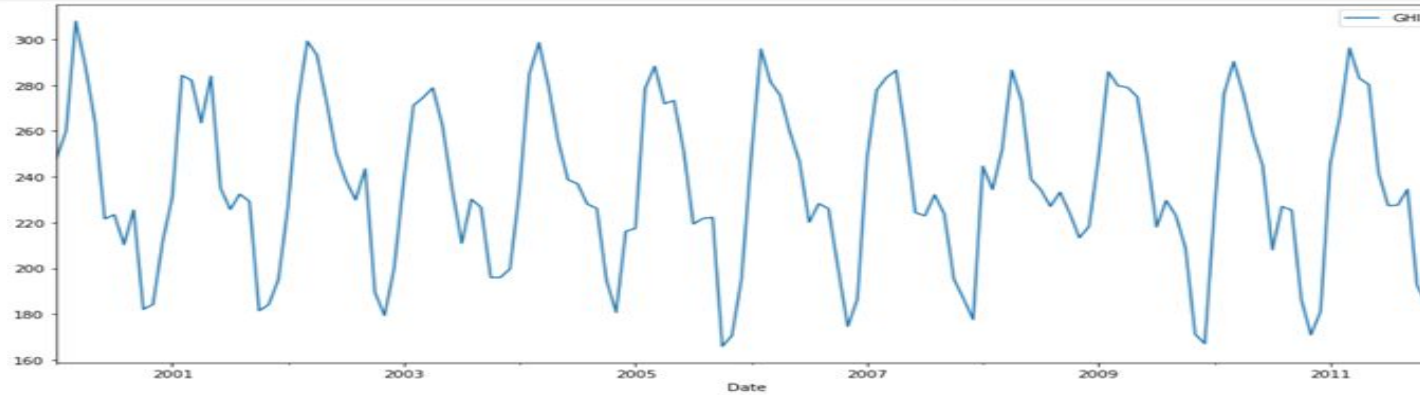




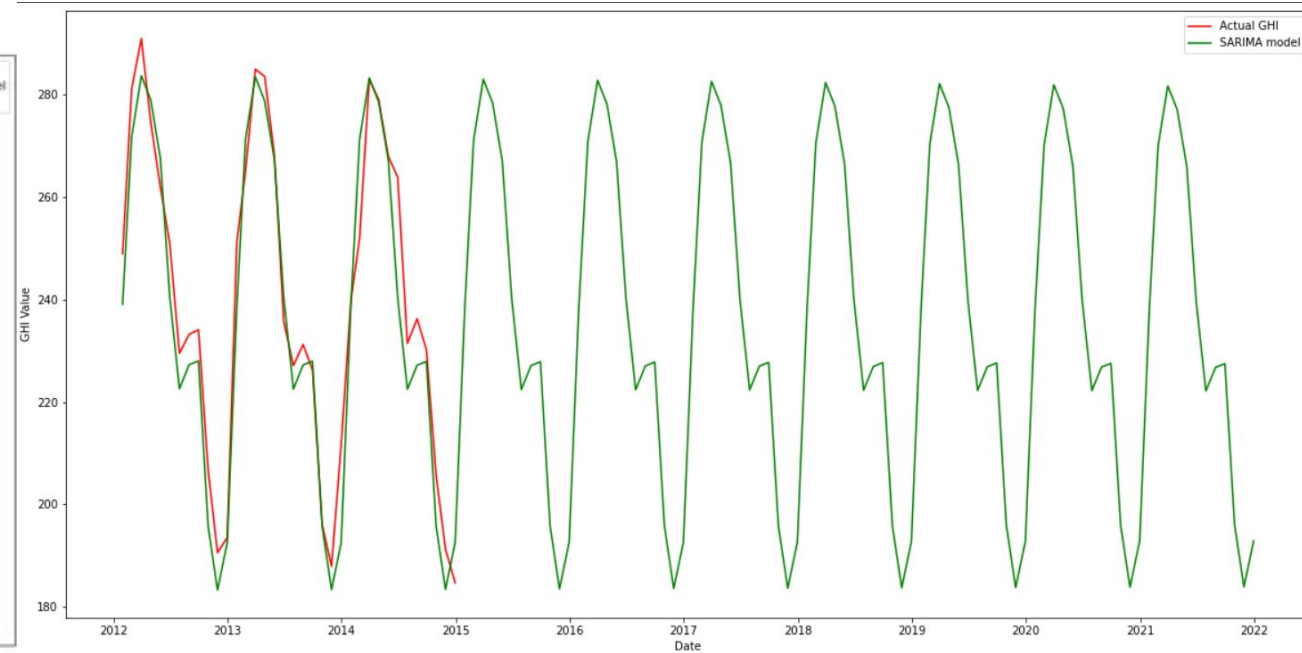
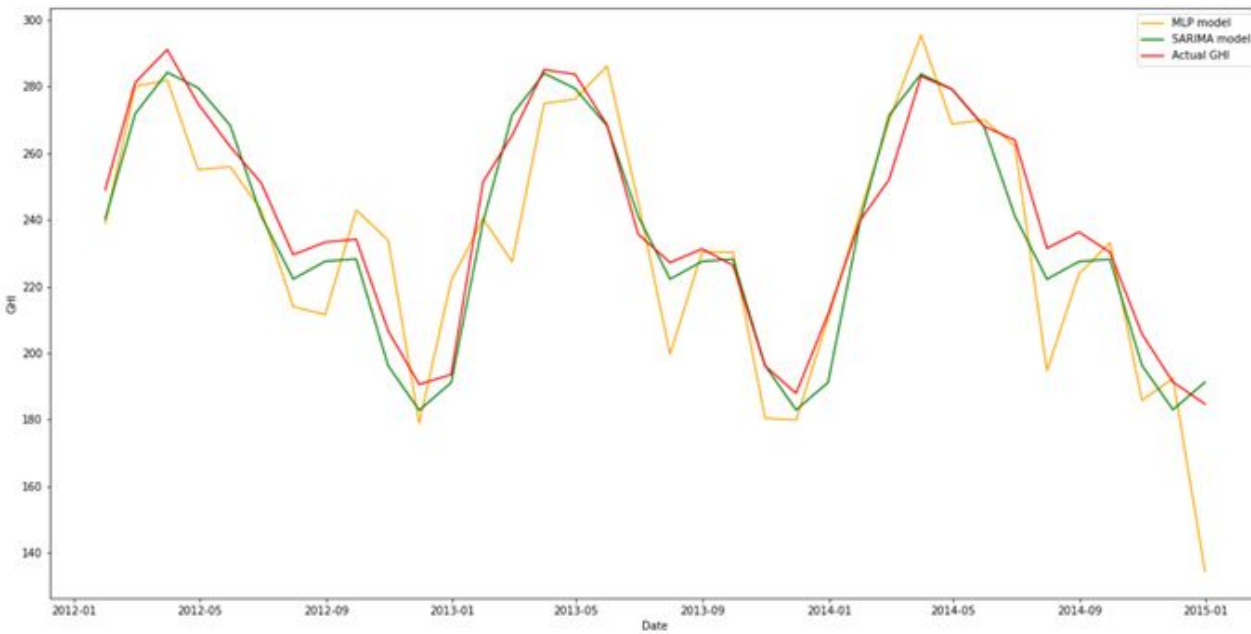
# Rajasthan (Weekly and Daily)



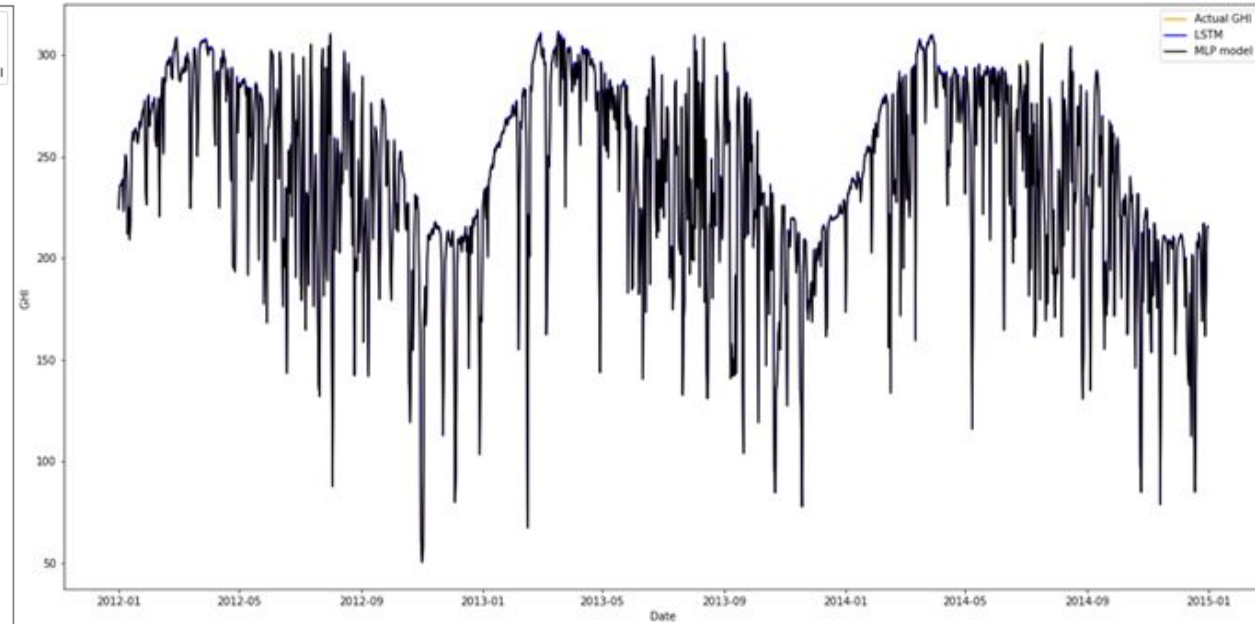
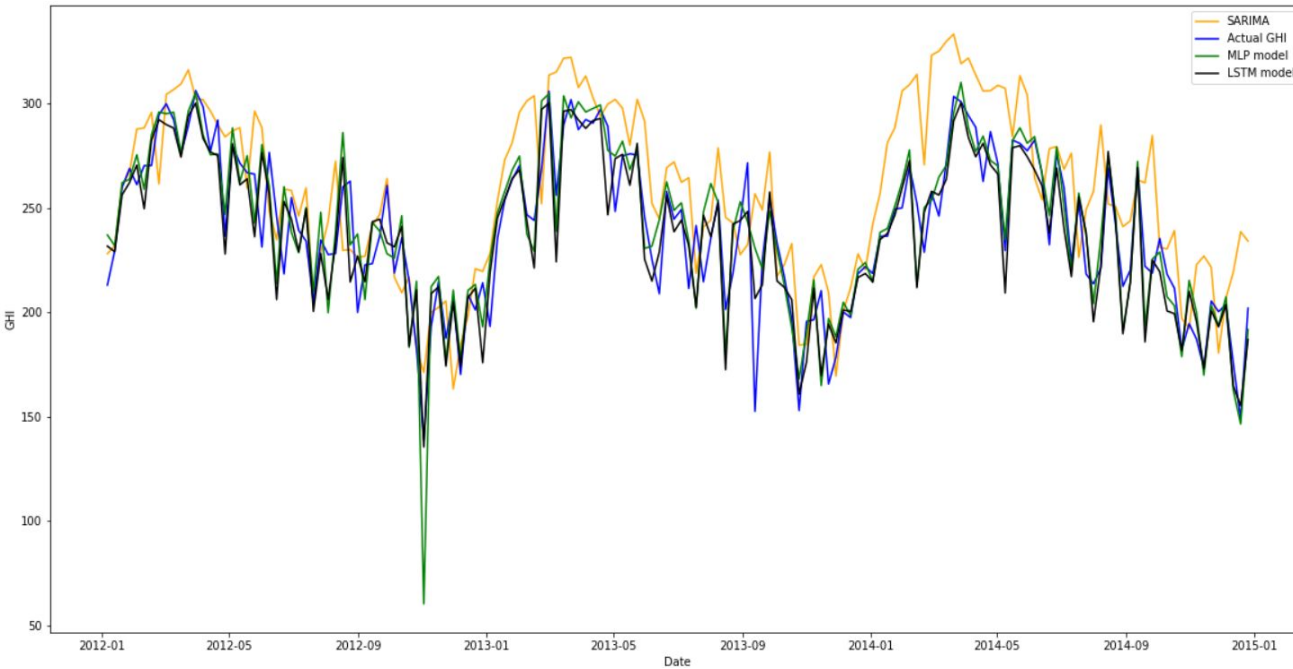
# Karnataka (Training Dataset)



# Karnataka(Monthly Predictions)

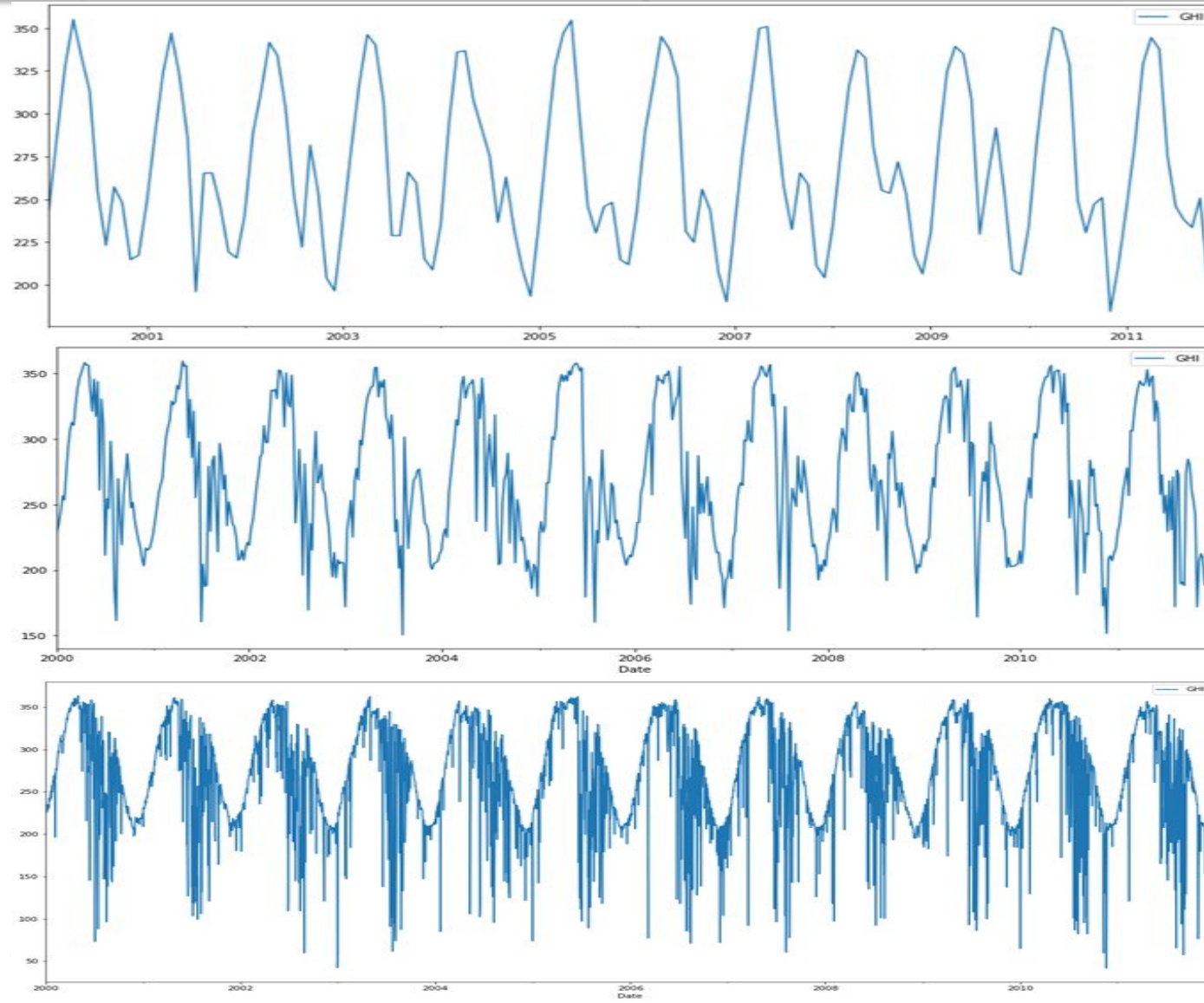


# Karnataka(Weekly and Daily)

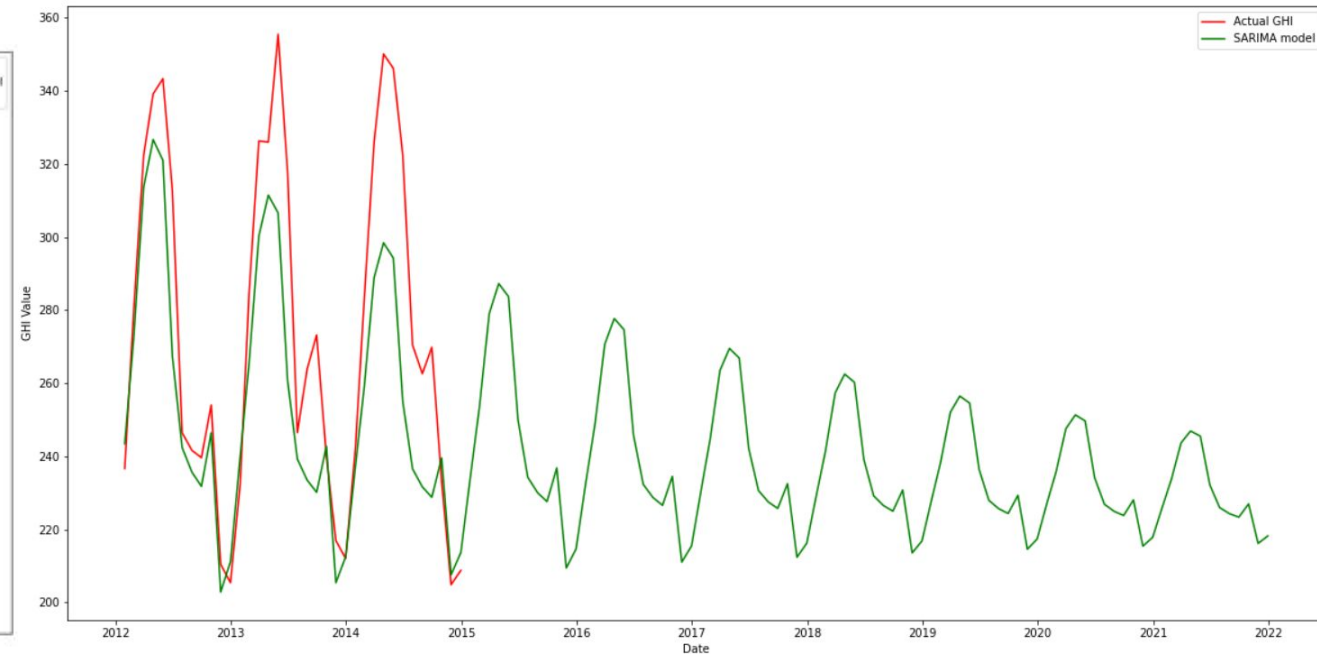




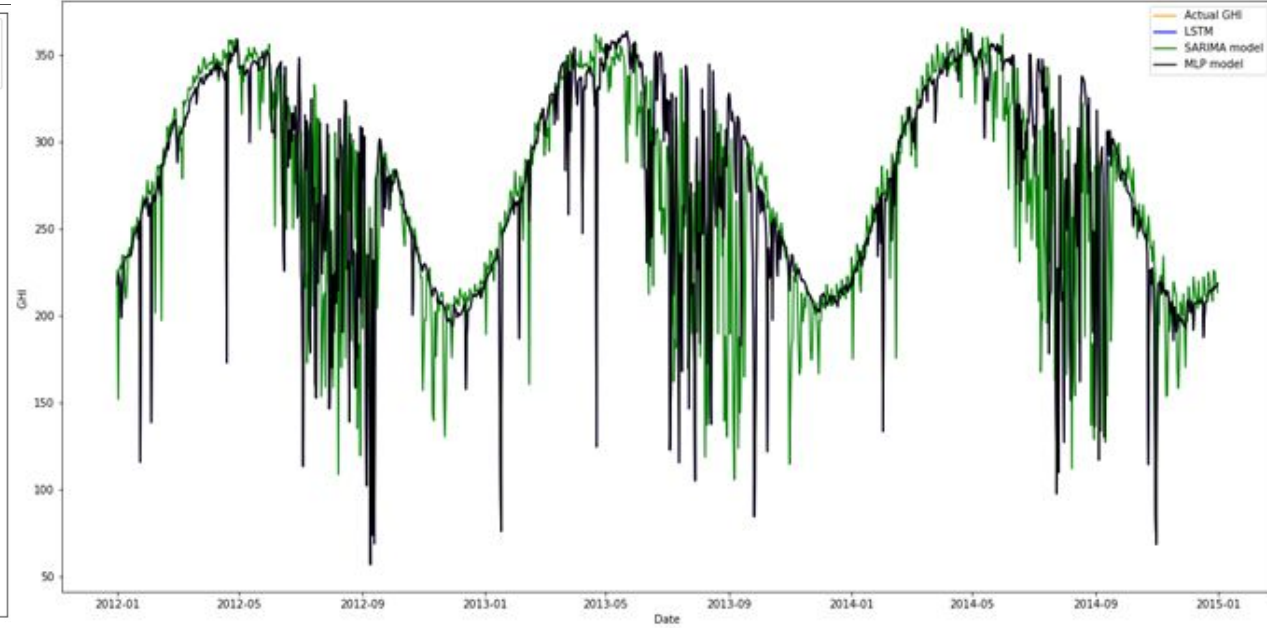
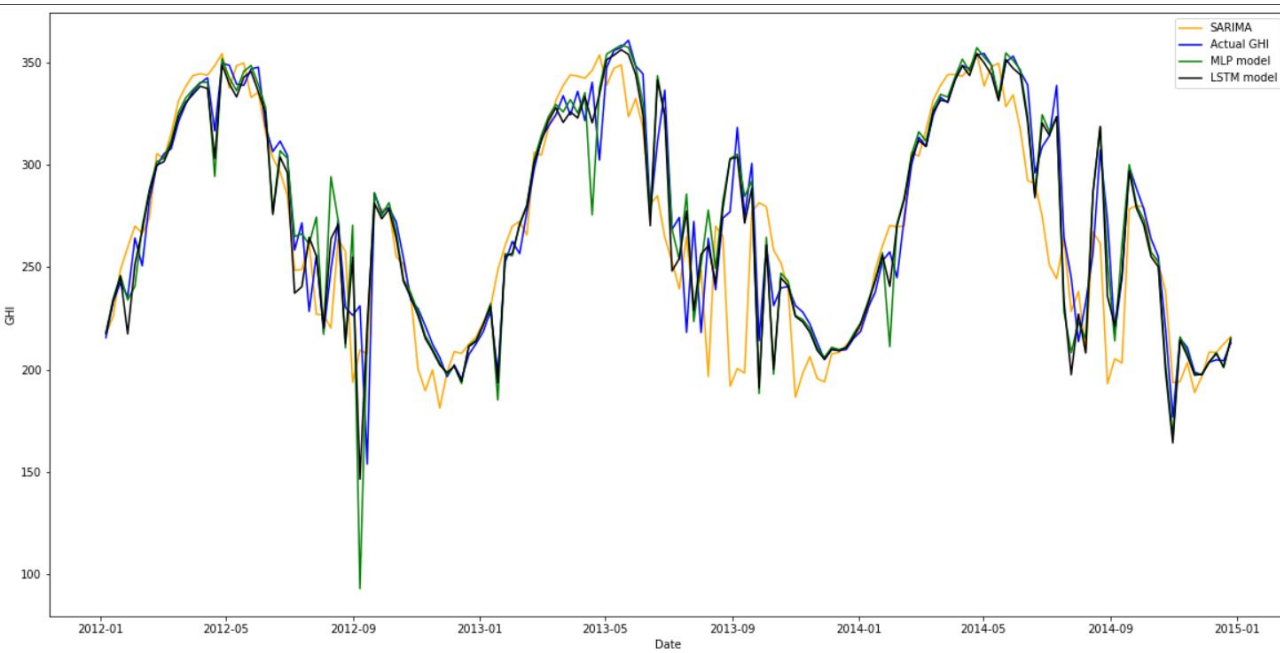
# Gujarat (Training Dataset)



# Gujarat (Monthly Predictions)

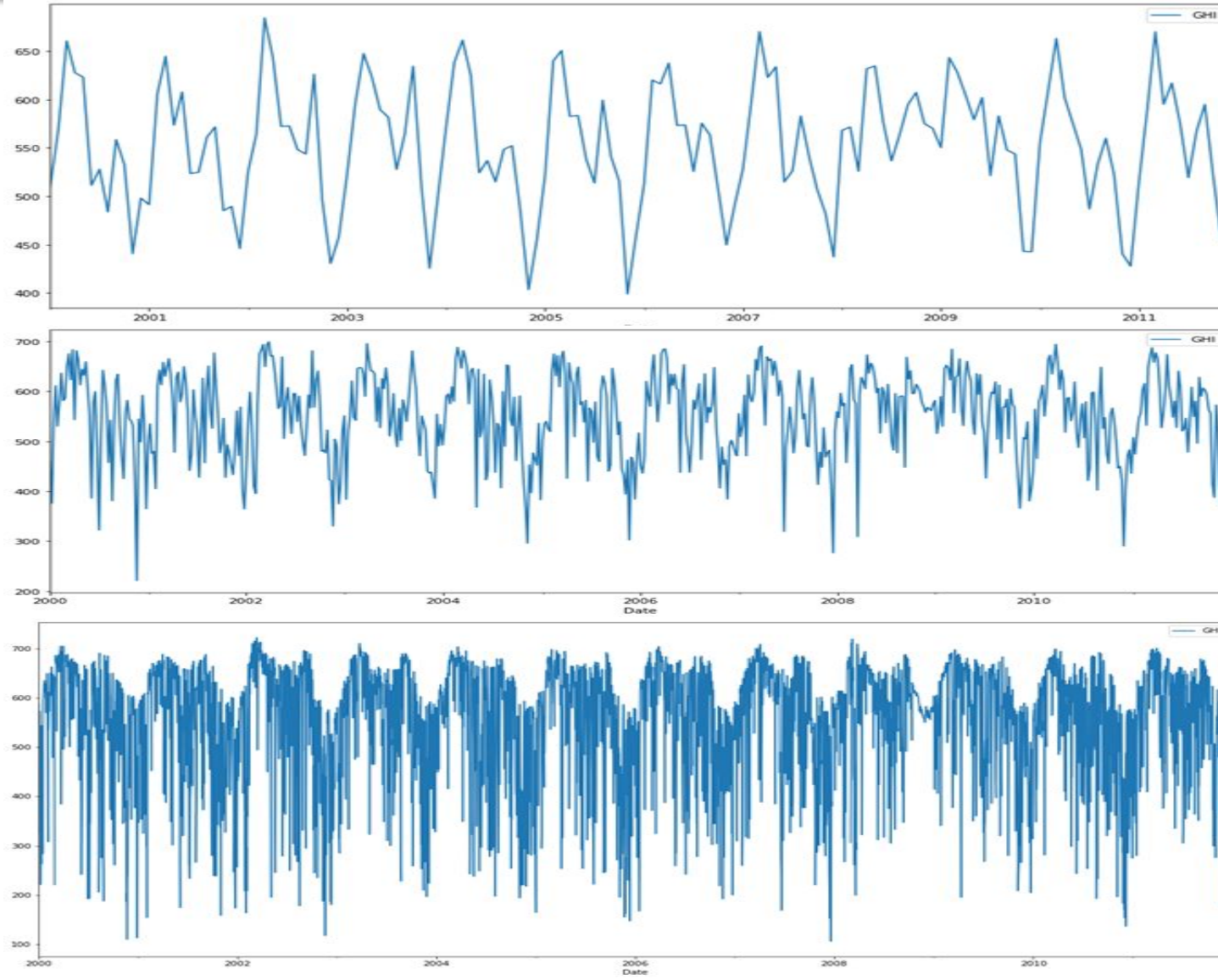


# Gujarat (Weekly and Daily)

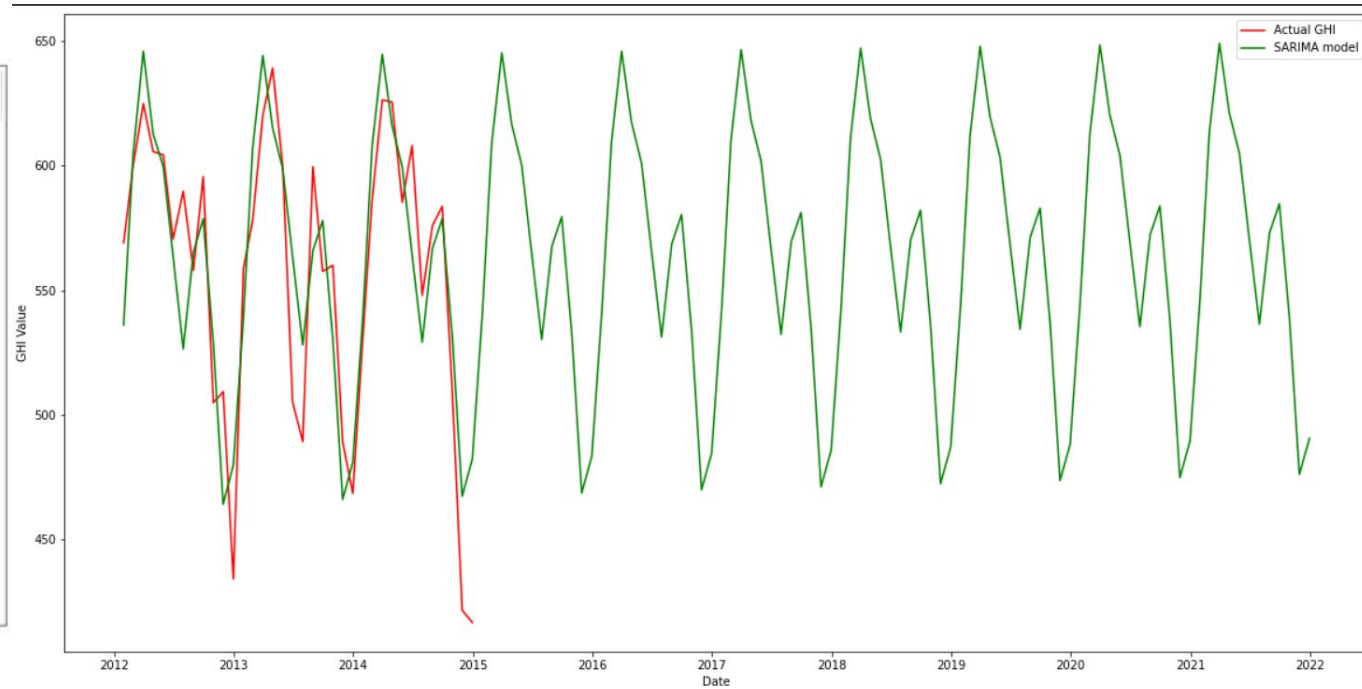




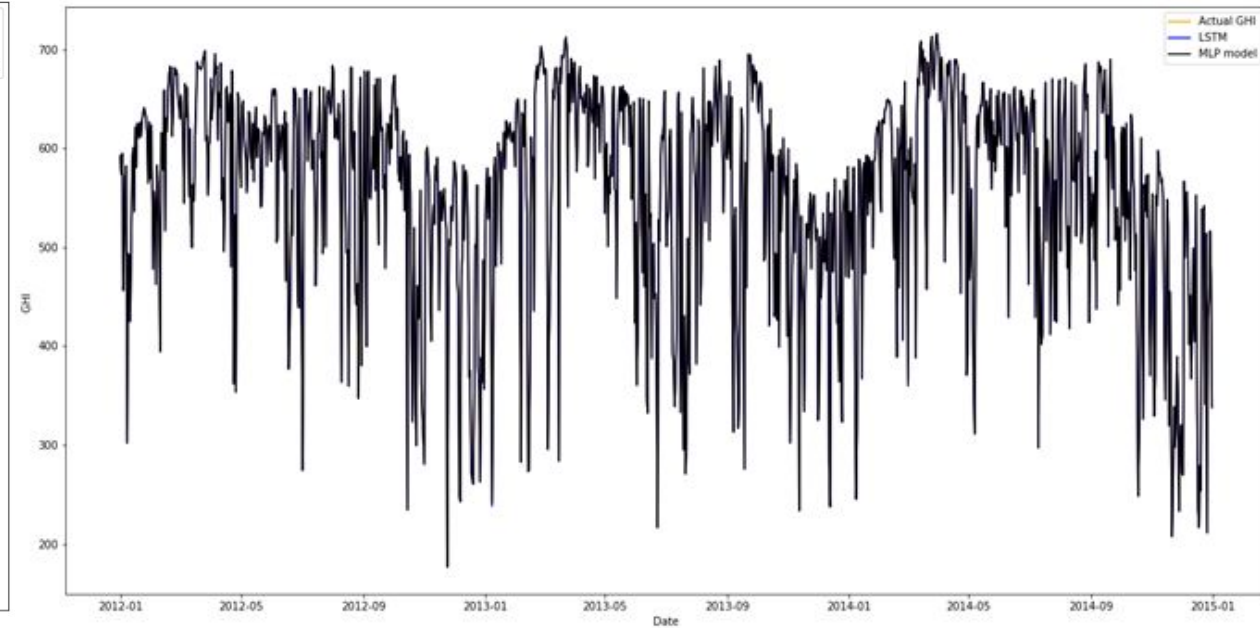
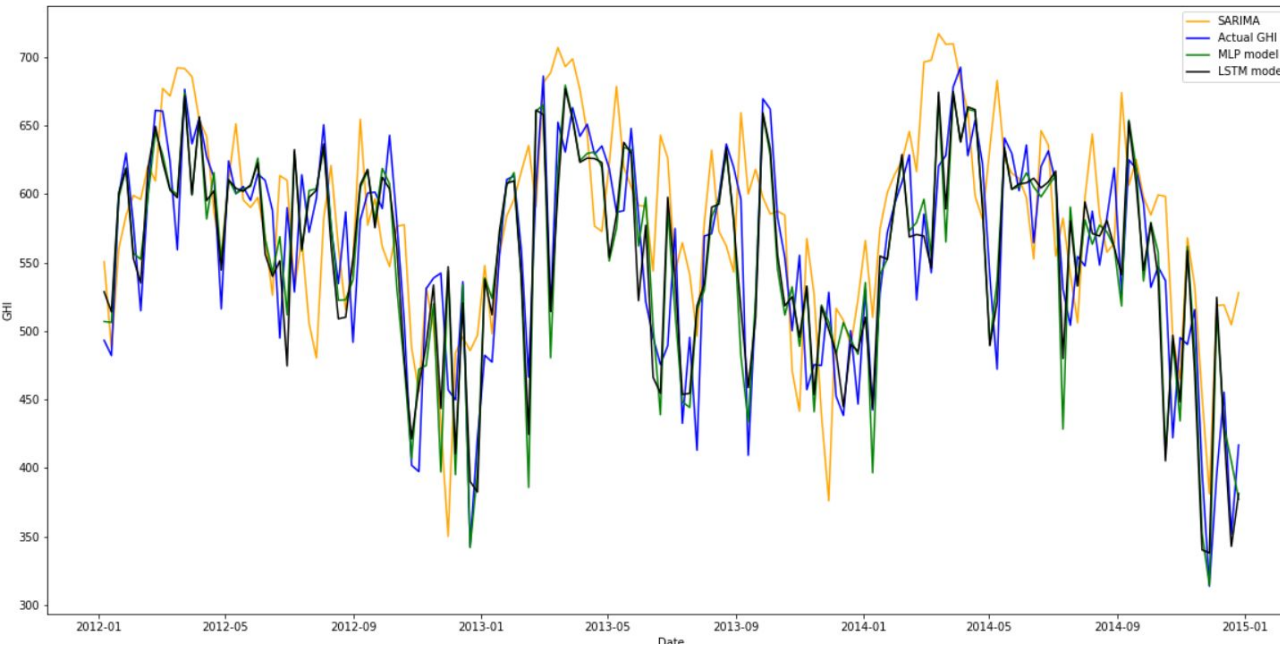
# Tamil Nadu(Training Dataset)



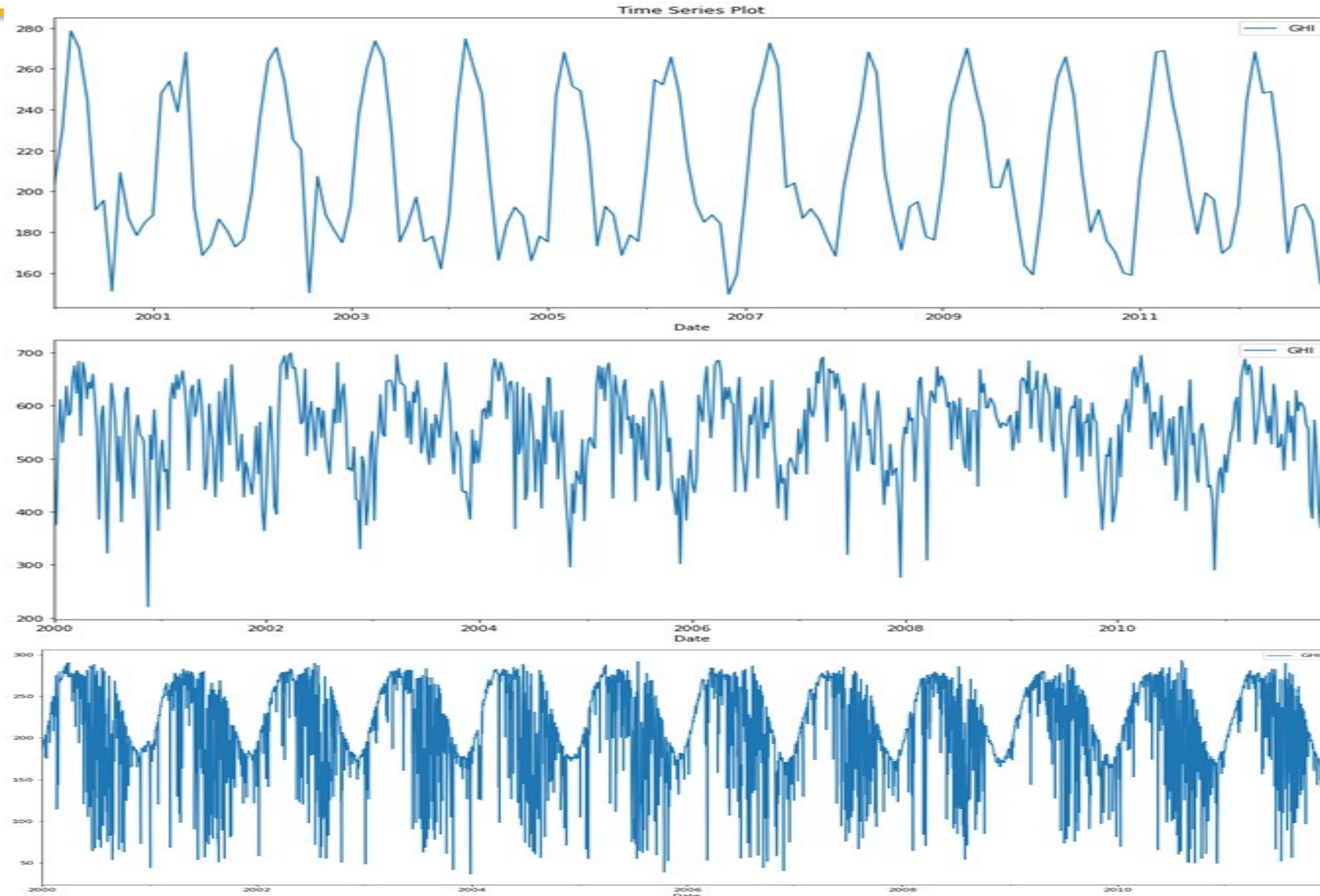
# Tamil Nadu(Monthly Predictions)



# Tamil Nadu(Weekly and Daily)

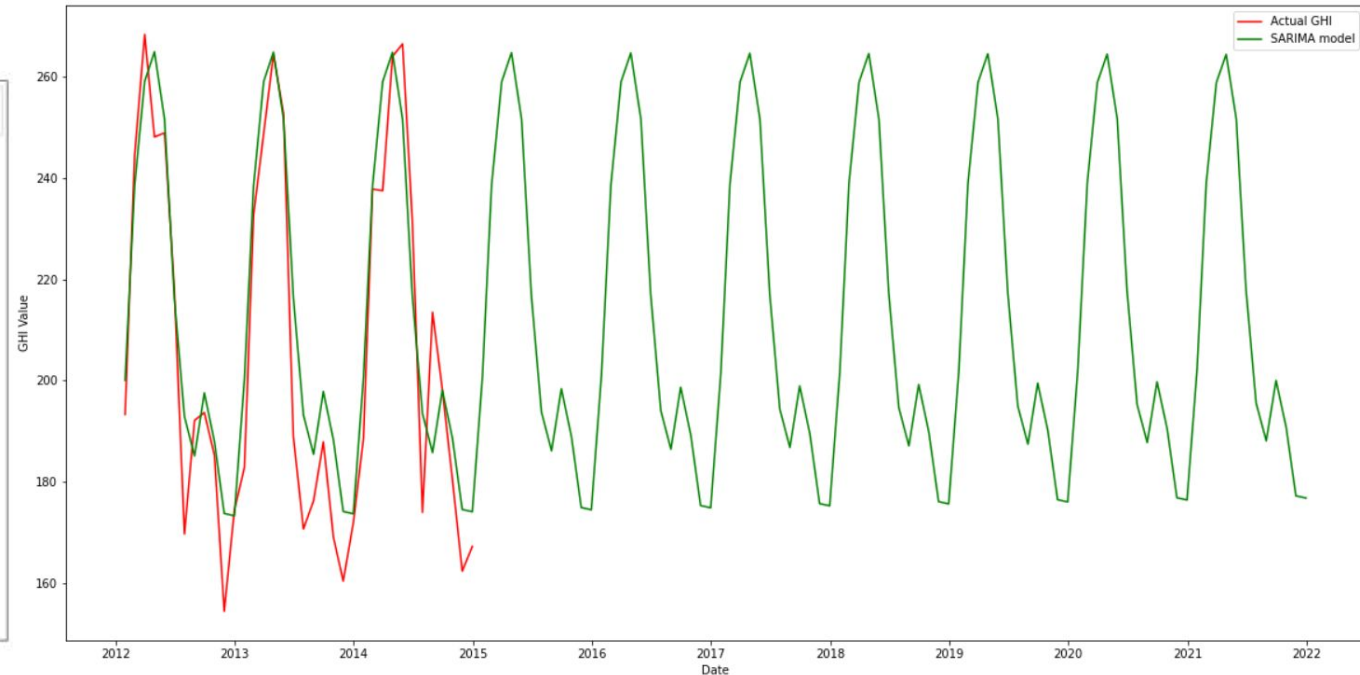


# Telangana(Training Dataset)

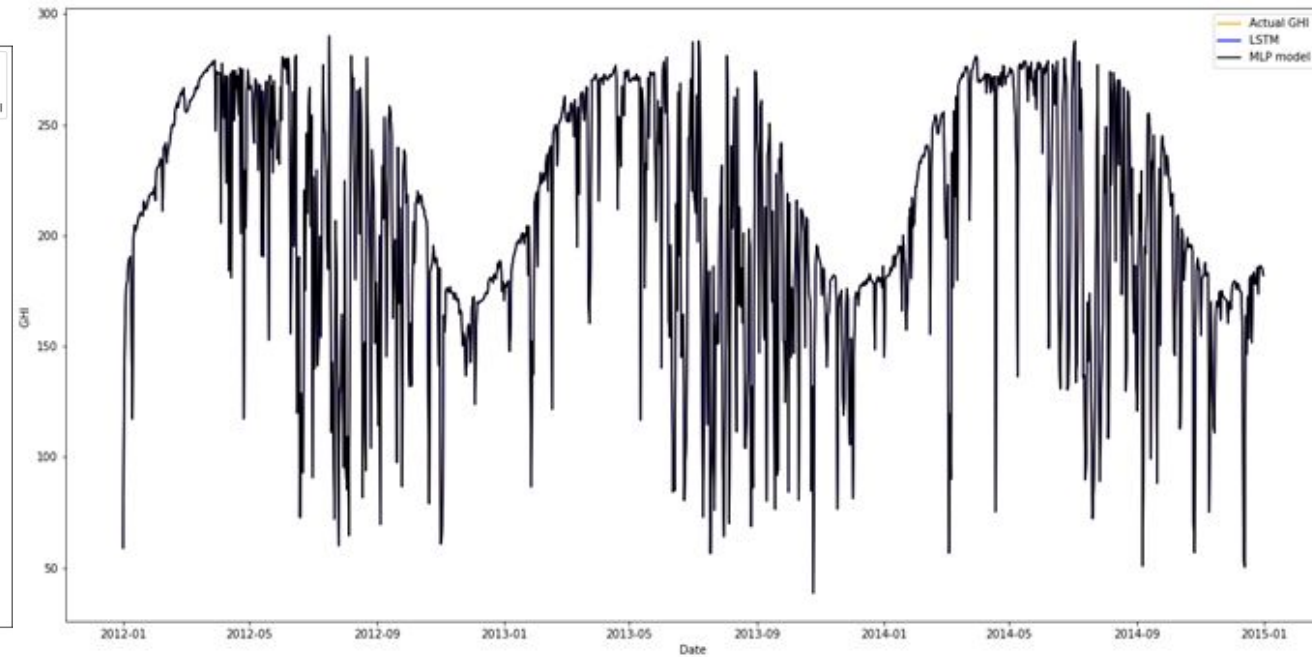
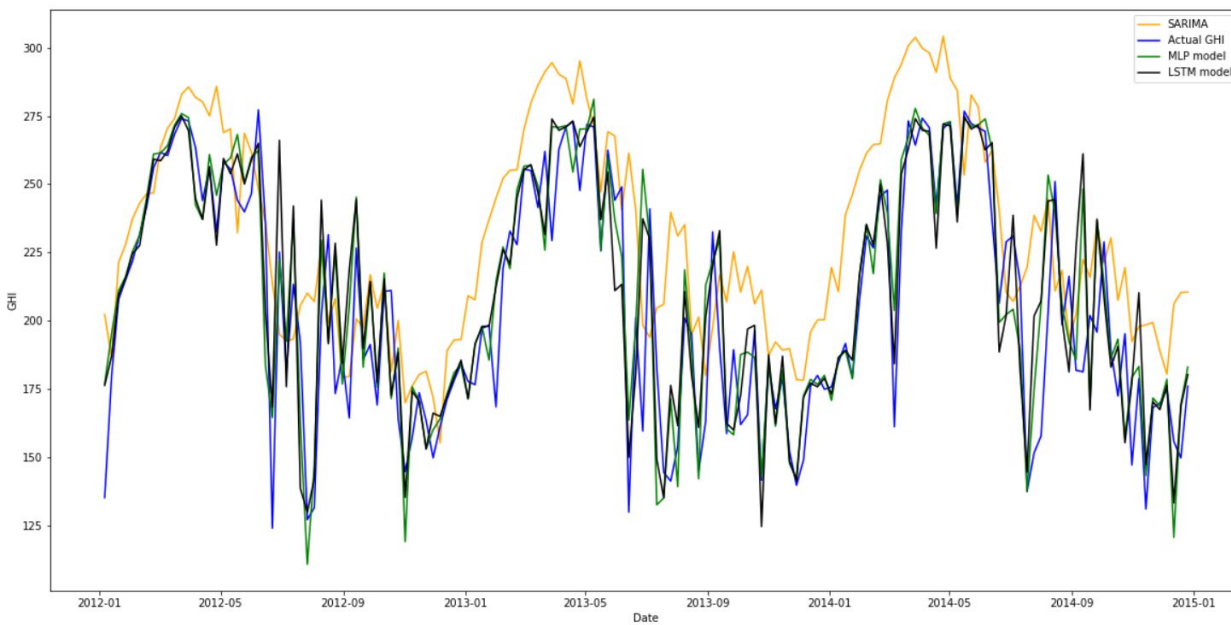




# Telangana(Monthly Predictions)



# Telangana(Weekly and Daily)



# Monthly Dataset



	MONTHLY DATASET		
STATE	MODEL	MAPE	RMSE
KARNATAKA	SARIMA(1,1,1)(1,0,1)[12]	2.83	8.57
ANDHRA PRADESH	SARIMA(1,1,0)(1,0,1)[12]	2.53	16.82
RAJASTHAN	SARIMA(1,1,1)(1,0,1)[12]	2.1	14.713
GUJARAT	MLP	4.12	21.24
TAMIL NADU	SARIMA(1,1,1)(1,0,0)[12]	4.59	29.46
TELANGANA	SARIMA(1,1,1)(1,0,1)[12]	5.45	13.28



# Weekly Dataset



	WEEKLY DATASET		
STATE	MODEL	MAPE	RMSE
KARNATAKA	LSTM	5.78	17.14
ANDHRA PRADESH	LSTM	4.60	35.089
RAJASTHAN	LSTM	2.98	23.99
GUJARAT	LSTM	4.06	16.12
TAMIL NADU	LSTM	6.61	43.74
TELANGANA	LSTM	8.54	22.4

# Daily Dataset



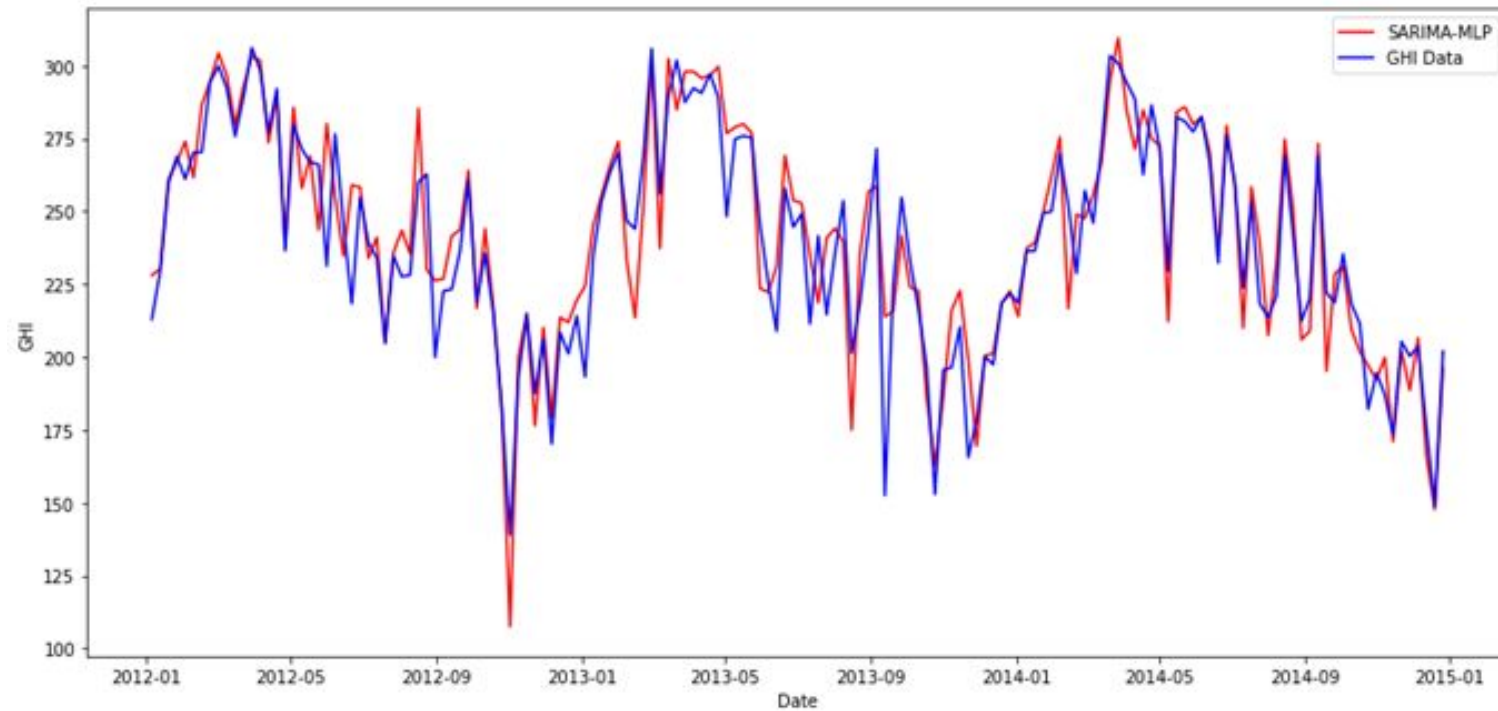
	DAILY DATASET		
STATE	MODEL	MAPE	RMSE
KARNATAKA	MLP, LSTM	0.16	0.399,0.443
ANDHRA PRADESH	MLP	0.0054	0.0741
RAJASTHAN	MLP	0.0097	0.078
GUJARAT	LSTM,MLP	0.015	0.0444,0.045
TAMIL NADU	LSTM	0.011	0.0718
TELANGANA	LSTM	0.017	0.036

# SARIMA-MLP Hybrid Models



## KARNATAKA

The SARIMA-MLP model for Karnataka dataset has MAPE value of 4.65 per cent and RMSE value of 14.44 which performs better than the LSTM model.

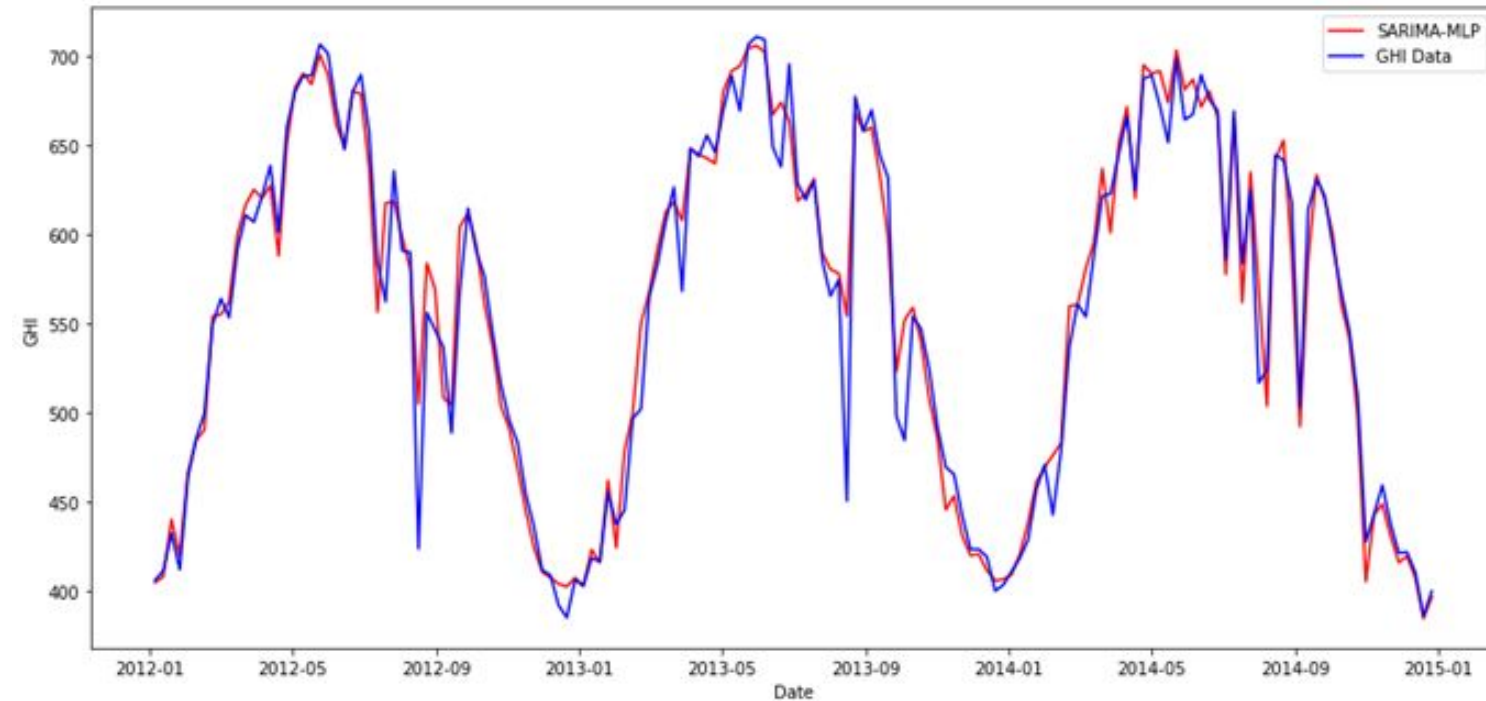


# SARIMA-MLP Hybrid Models



## ANDHRA PRADESH

The SARIMA-MLP model for Andhra Pradesh dataset has MAPE value of 2.31 per cent and RMSE value of 19.25 which performs better than the LSTM model.

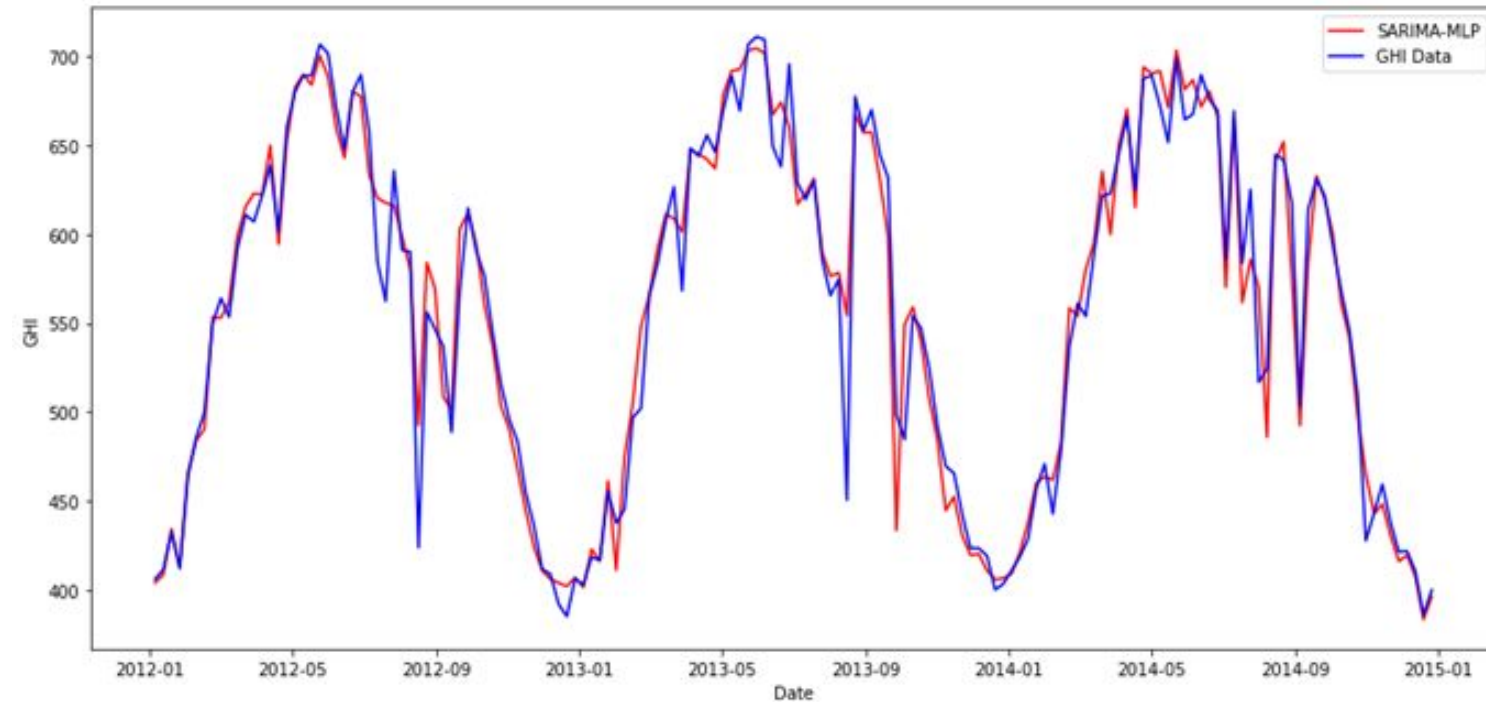


# SARIMA-MLP Hybrid Models



## RAJASTHAN

The SARIMA-MLP model for Rajasthan dataset has MAPE value of 2.47 per cent and RMSE value of 20.21 which performs better than the LSTM model.

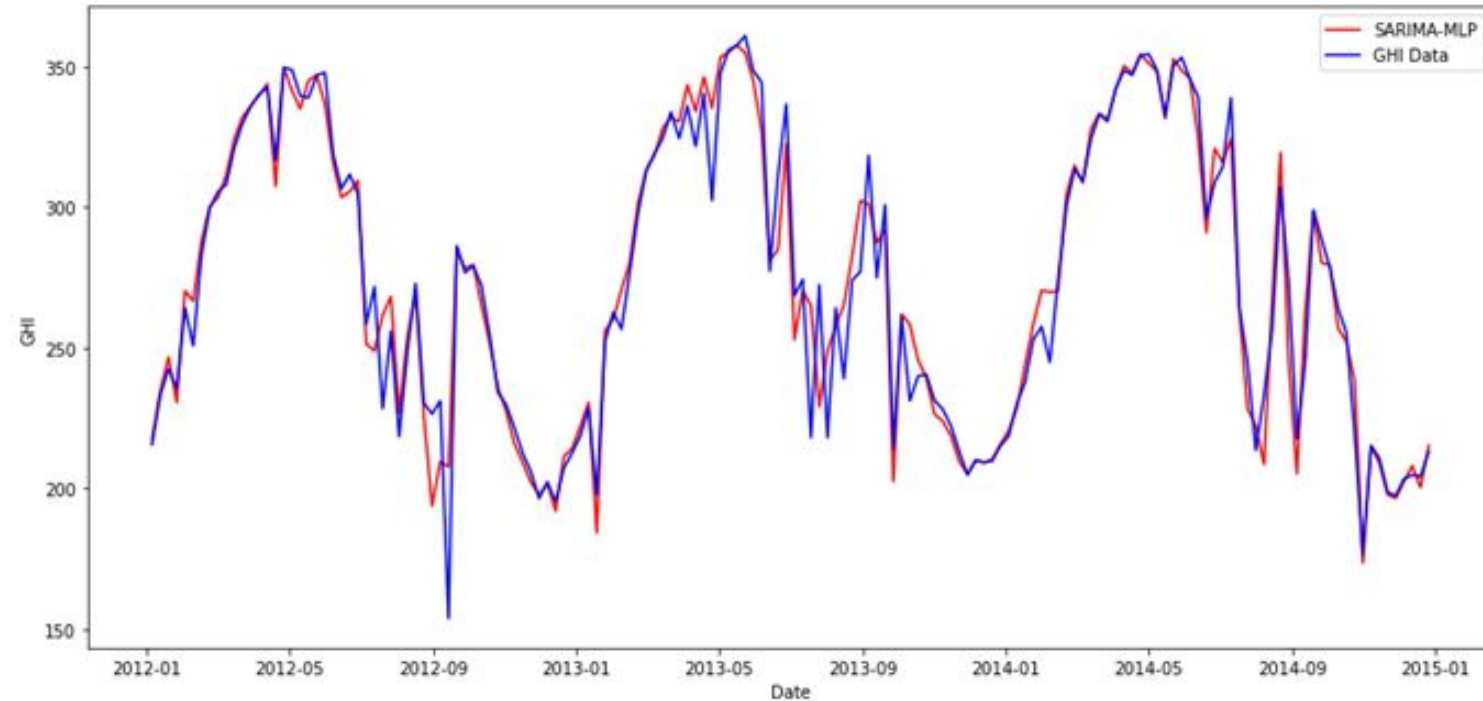


# SARIMA-MLP Hybrid Models



## GUJARAT

The SARIMA-MLP model for Gujarat dataset has MAPE value of 2.99 per cent and RMSE value of 12.19 which performs better than the LSTM model.

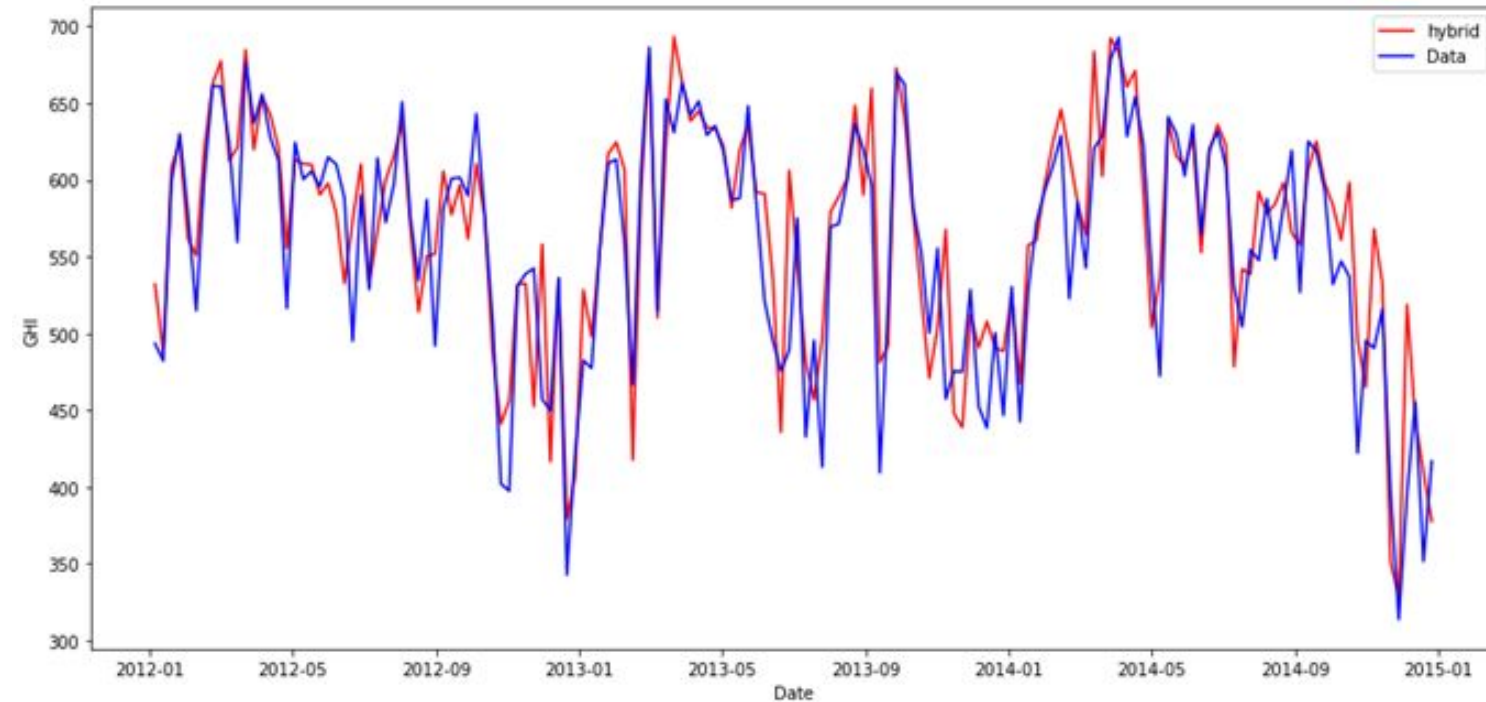


# SARIMA-MLP Hybrid Models



## TAMIL NADU

The SARIMA-MLP model for Tamil Nadu dataset has MAPE value of 5.53 per cent and RMSE value of 31.92 which performs better than the LSTM model.



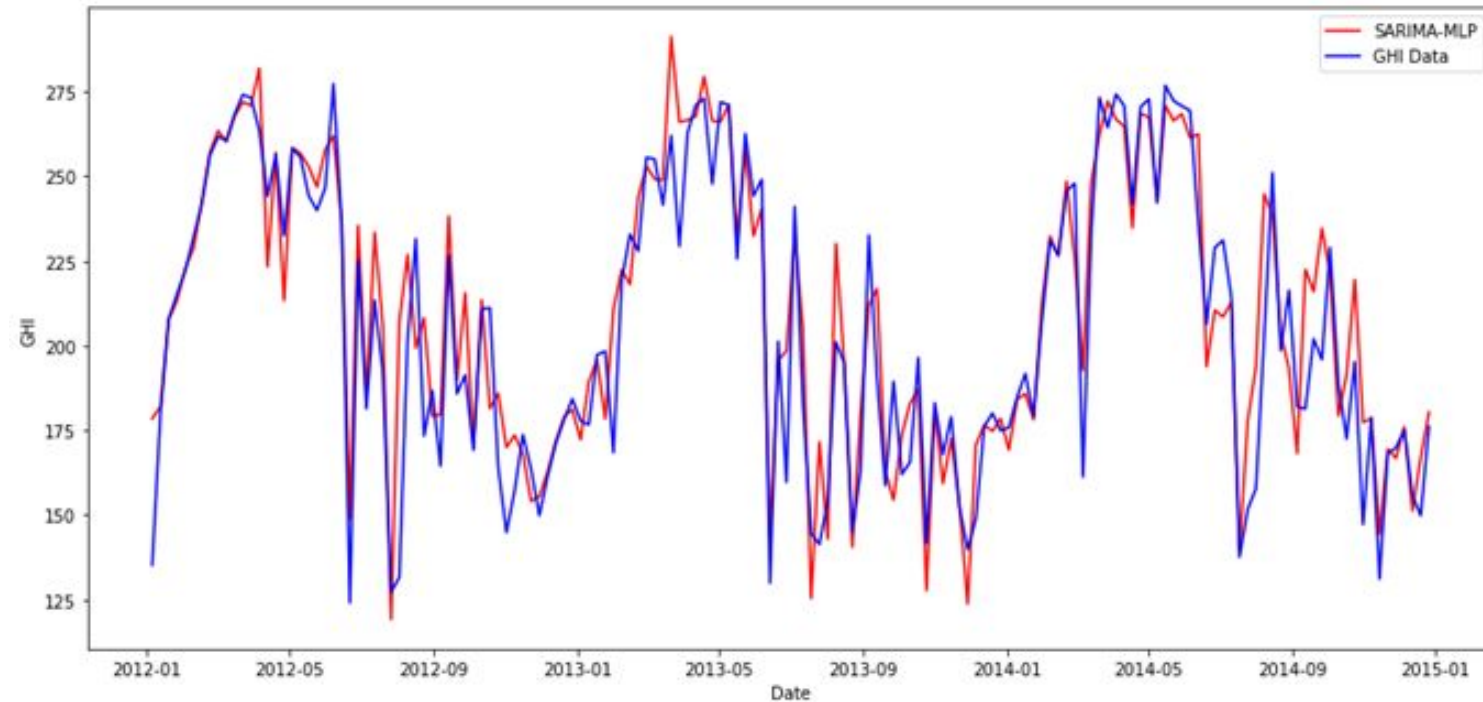


# SARIMA-MLP Hybrid Models



## TELANGANA

The SARIMA-MLP model for Telangana dataset has MAPE value of 6.59 per cent and RMSE value of 17.19 which performs better than the LSTM model.



# Observations and Conclusions



1. GHI index is highly correlated to DHI, DNI, Clearsky GHI, Clearsky DNI, Clearsky DHI and Temperature.
2. From the time series decomposition, we can clearly see there is a clear seasonal component in GHI.
3. The solar sites situated in Rajasthan, Andhra Pradesh and Tamil Nadu had higher median GHI values than solar sites in Karnataka, Gujarat and Telangana.
4. Solar site situated in Andhra Pradesh had least coefficient of variance followed by Rajasthan, but the number of outliers in Andhra Pradesh dataset is greater than Rajasthan. Gujarat and Karnataka had almost identical coefficient of variance. Telangana had the greatest coefficient of variance.
5. The SARIMA model performs best for the monthly dataset for all states except Gujarat which MLP model gave best predictions. The most accurate results were obtained for solar site located in Rajasthan.

# Observations and Conclusions



6. For the weekly dataset, LSTM models outperform all the other models. Again the most accurate predictions were obtained for Rajasthan.
7. For the daily dataset, the predictions of LSTM and MLP models give similar results. Solar sites in Andhra Pradesh and Rajasthan had most accurate results from MLP model, while Tamil Nadu and Telangana favoured LSTM model. Karnataka and Gujarat had similar results from LSTM and MLP models. The most accurate results were obtained from the solar site in Andhra Pradesh.
8. The SARIMA-MLP hybrid models outperformed LSTM model for weekly dataset of all the states.
9. Conducting visual analysis on the SARIMA-MLP model of Rajasthan shows that majority of the occurrences where SARIMA predictions are preferred over MLP predictions occurred during the Monsoon season of Rajasthan.
10. The MLP models had the least training times in comparison to the other models while the LSTM model had the greatest training time for the above datasets.

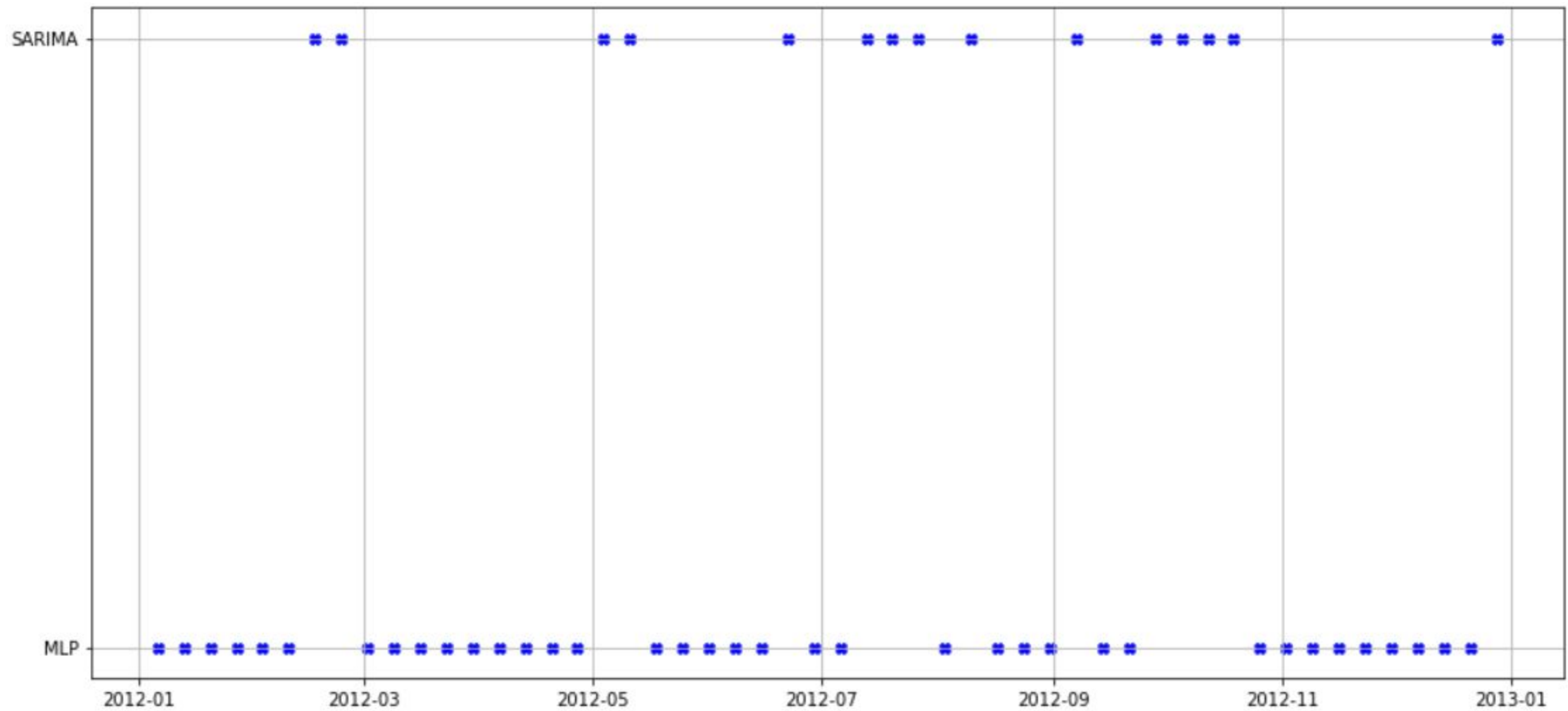
# References



- National electricity plan (2016), Volume 1, Generation, Central Electricity Authority (CEA), Ministry of Power, GOI .Available at [http://www.cea.nic.in/reports/committee/nep/nep\\_dec.pdf](http://www.cea.nic.in/reports/committee/nep/nep_dec.pdf) .
- Reikard, G. (2009). Predicting solar radiation at high resolutions: A comparison of time series forecasts. Solar Energy, 83(3), 342-349.
- Al-Sadah, F. H., Ragab, F. M., & Arshad, M. K. (1990). Hourly solar radiation over Bahrain. Energy, 15(5), 395-402.
- Wu Ji, Keong Chan Chee, Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN, Solar Energy, Volume 85, Issue 5, May 2011, Pages 808-817.
- M. Bouzardoum, A. Mellit, A. Massi Pavan, A hybrid model (SARIMA-SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant, Sol. Energy 98 (PC) (2013) 226e235
- Jie Shi, Wei-Jen Lee, Yongqian Liu, Yongping Yang, and Peng Wang, "Forecasting Power Output of Photovoltaic Systems Based on Weather Classification and Support Vector Machines," IEEE Transactions on Industry Applications, vol. 48, no. 3, pp. 1064–1069, 2012.
- Wu, Y.-K., Chen, C.-R., and Abdul-Rahman, H. A novel hybrid model for short-term forecasting in PV power generation. International Journal of Photoenergy, 2014(0):1 –9, 2014. doi: <http://dx.doi.org/10.1155/2014/569249>.
- Power Sector at a Glance All India. (2020, 02 02). Retrieved from Ministry of Power Gol: <https://powermin.nic.in/en/content/power-sector-glance-all-india>
- *Brown to Green: The G20 Transition Towards A Net-Zero Emissions Economy 2019 report*. (2020, 02 02). Retrieved from Enerdata 2019: <https://www.climate-transparency.org/wp-content/uploads/2019/11/Brown-to-Green-Report-2019.pdf>
- Each Country's Share of CO2 Emissions. (2020, 02 02). Retrieved from Union of Concerned Scientists: <https://www.ucsusa.org/resources/each-countrys-share-co2-emissions>
- World Energy Scenarios Composing energy futures to 2050 (2013), World energy Council. [https://www.worldenergy.org/wp-content/uploads/2013/09/World-Energy-Scenarios\\_Composing-energy-futures-to-2050\\_Full-report.pdf](https://www.worldenergy.org/wp-content/uploads/2013/09/World-Energy-Scenarios_Composing-energy-futures-to-2050_Full-report.pdf) .
- India's potential for integrating solar and on- and offshore wind power into its energy system. (2020, 02 02). Retrieved from Nature Communications: <https://www.nature.com/articles/s41467-020-18318-7>
- Annual Report Ministry of New and Renewable Energy 2019. (2020, 02 05). Retrieved from Ministry of New and Renewable Energy, Gol: [https://mnre.gov.in/img/documents/uploads/file\\_f-1597797108502.pdf](https://mnre.gov.in/img/documents/uploads/file_f-1597797108502.pdf)

Thank  
You

# APPENDIX





# APPENDIX

innovate

achieve

lead

NEWS

• LIVE TV

INDIA  
TODAY

APP

MAGAZINE

HOME

MY FEED

VIDEOS

CORONA

ELECTIONS

IPL

MALAYALAM

INDIA

FACT CHECK

QUIZ

EDUCATION

□ □ □



News / India / Monsoon likely to arrive in Rajasthan on June 25: MeT

## Monsoon likely to arrive in Rajasthan on June 25: MeT

*The meteorological department said south-west monsoon normally enters the desert state from its southern region and its normal time of arrival is June 15. It continues till September 20,*



Press Trust of India

New Delhi

May 14, 2020 UPDATED: May 14, 2020 04:19 IST



### READ THIS

Ambulance drivers in Delhi risk it all to ferry Covid-19 victims



Covid second wave: The cost of lockdowns



How campaign spiked Covid fatalities, infections in election states



Covid second wave: All your FAQs answered at doctors' roundtable



### RECOMMENDED

# Need for Forecasting



When the traditional sources like fossil fuels are used for power generation, the generation is predominantly controlled by generation capacity of the plant.

But when renewable sources like Solar and Wind energy are used for power generation, the generation also depends on weather conditions apart from the capacity of the machines.

Thus, there is a need to build forecast models for generation in order to have a better generation scheduling.

This study project will primarily focus on Solar Energy. Considering the advantages of Solar energy over other non-renewable sources, the development and research on solar power has been rising year by year.