COBEM 2023 27TH INTERNATIONAL CONGRESS OF MECHANICAL ENGINEERING

ABCM Brazilian Society of Mechanical Sciences and Engineering
27th ABCM International Congress of Mechanical Engineering
December 04-08, 2023, Florianópolis-SC, Brazil

COB-2023-1621
PREDICTING SOLAR RADIATION IN MINAS GERAIS USING ARTIFICIAL INTELLIGENCE TECHNIQUES

João Vitor Silva Nazareth
Ricardo Henrique Guedes Furiati
Silvio Jamil Ferzoli Guimarães
Felipe Domingos da Cunha
Cristiana Brasil Maia
Pontifícia Universidade Católica de Minas Gerais
nazareth@sga.pucminas.br, rfuriati@sga.pucminas.br, sjamil@pucminas.br, felipe@pucminas.br, cristiana@pucminas.br

Abstract. *Approximately 80% of the world´s energy is generated from non-renewable sources. Fossil fuels are depleting and will not be able to meet the global energy demand in a near future. Solar energy is one of the cleanest sources and it is seen as one of the main resources for massive renewable energy generation. It is important to provide a deeper understanding of solar energy resources to successfully implement solar energy projects. Therefore, accurate prediction of solar radiation is essential to improve its use. In this paper, three machine learning methods (SARIMA, Holt-Winters, and LSTM) are used to predict the global solar radiation in the state of Minas Gerais, Brazil. The results were compared to real data collected from NASA satellites. For the city of Belo Horizonte, the results were also compared to results from an analytical model. Holt-Winters method presented the best performance, with a maximum difference of approximately 5.0% compared to real data. This method was used to generate a monthly solar radiation map for the state of Minas Gerais, using data from 90 cities in the state and vicinities.*

Keywords: *Solar Radiation, Machine Learning, SARIMA, Holt-Winters, LSTM, Solar Radiation Map*

1. INTRODUCTION

In 2017, the world's total primary energy supply was 13972 Mtoe, with coal and oil corresponding to, respectively, 27.1 and 32.0% (Staff, 2019). Fossil fuels are, nevertheless, nonrenewable and are depleting faster than new reserves can be created. It is important to use renewable energy sources to fulfill the world's energy demands. This trend has been observed in the last years, with an increase in the share of renewable energy sources in primary energy consumption. According to (Yolcan, 2023), while 0.24 Gtoe of the world's primary energy consumption was met from renewable energy sources in 2012, this value increased to 0.95 Gtoe in 2021. In proportional terms, the share of renewables rose from 2.0% in 2012 to 6.7% in 2021.

Brazil ranks third globally in terms of cumulative energy capacity with 150 GW, following China (908 GW) and the United States (313 GW). In 2020, the country ranked third globally in wind power additions, adding three times more wind power capacity than in 2019. Renewable energy capacity investments in Brazil grew by 23% in 2020, marking the seventh consecutive year of growth, accounting for 2.9% of total global renewable energy capacity investments in 2020. It is a world leader in the production and use of bioenergy, holding a 25% share of the global production of biofuels. Nevertheless, its production relies mainly on hydropower (REN21, 2021). While ranking third globally in terms of cumulative energy capacity when excluding hydropower, the ranking in renewable power capacity per person is only 41. It is important to diversify the energy matrix.

When comparing renewable energy sources, solar energy is the most potentially sustainable renewable energy source (Kumar *et al.*, 2023).It is now the cheapest and most competitive source of new electricity generation in most markets worldwide (Formolli *et al.*, 2023) . It has the potential to meet energy demands in terms of sustainability and quality. The solar energy that falls on Earth's continents is more than 200 times greater than the annual total commercial power currently consumed by humans (Kumar *et al.*, 2023). Solar energy harvesting techniques can be broadly classified into two categories: (1) direct electricity generation using solar photovoltaic panels; (2)indirect conversion using solar thermal collectors (Pandey *et al.*, 2022).The proper design of solar equipment depends on the knowledge of the availability of solar energy.

Solar radiation measurements are highly important for achieving energy efficiency in smart buildings as well as solar energy production. They are commonly acquired with pyranometer sensor devices. However, due to its high initial and maintenance costs, it is not densely deployed in the field. Consequently, it provides only limited coverage as a data source for solar radiation. Hence, theoretical, empirical, and/or data-driven models are utilized to estimate solar radiation in areas without pyranometers (Kosovic *et al.*, 2020). Solar radiation value depends on many parameters such as sun angle, hour angle, topography, air composition, morphology area, and water as well as climate conditions. Over the years, various mathematical models have been developed to estimate worldwide solar radiation. Empirical models have been widely studied with the correlation between weather characteristics and the position of the sun (Budiyanto and Lubis, 2020).

Although a significant number of empirical mathematical models are established in the literature, their ability to model the sun's activity is directly tied to the complexity of each model and the number of variables used. Furthermore, the direct application of empirical models is limited to the understanding of complex mathematical equations and intricate variables not readily available. To mitigate this problem, using artificial intelligence methods can significantly help in modeling the sun's future activity more efficiently.

To that end, this work presents three forecasting models- LSTM, SARIMA and Holt-Winters- used to estimate future solar radiation values for the Brazilian state of Minas Gerais. The selected models present a flexible approach to predict solar behavior, given that all three models use only one variable, analyzing historical past solar irradiance data to predict future values.

The data used for this study comprises the most expansive dataset found in the literature, with 35 years of solar irradiance data collected by NASA satellites used to predict future values for 90 locations in the State of Minas Gerais and its vicinities. For the city of Belo Horizonte, each model's results were compared to the results of an analytical model. To better visualize the results of the models, the predictions were interpolated to generate a solar radiation map for the state.

2. MATHEMATICAL MODEL

It was estimated the total solar radiation incident on a horizontal surface, following the mathematical model presented in (Duffie and Beckman, 2013). Although the results of the machine learning methods have been used to predict the solar radiation in the state of Minas Gerais, the comparison with the analytical model was performed only for the city of Belo Horizonte (latitude 19.98 S and longitude 43.97 W).

The daily solar radiation H is obtained based on the daily extraterrestrial radiation on a horizontal surface, H_0, in J/m^2.

$$H_0 = \frac{24x3600Gsc}{\pi}\left(1 + 0.033cos\frac{360n}{365}\right) * \left(cos\phi\, cos\delta sin\omega_s + \frac{\pi\omega_s}{180}sin\phi sin\delta\right) \tag{1}$$

Gsc is the solar constant, or the energy from the sun per unit time received on a unit area of the surface perpendicular to the direction of propagation of the radiation at a mean earth-sun distance outside the atmosphere, assumed as 1367 W/m^2. n is the day of the year, varying from 1 to 365, ϕ s the latitude, δ is the declination, or the angular position of the sun at solar noon with respect to the plane of the equator, and ω_s is the sunset hour angle.

The declination is given by:

$$\delta = 23,45sin\left(360\frac{284+n}{365}\right) \tag{2}$$

The sunset hour angle is given by:

$$\omega_s = acos(-tg(\phi).tg(\delta)) \tag{3}$$

The daily solar radiation H is given by:

$$K_t = \frac{H}{H_0} \tag{4}$$

K_t is the daily clearness index, or the ratio of a particular day's radiation to the extraterrestrial radiation for that day. The value of K_t is not defined for a specific day for a specific site. It is required a series of historical data for its determination. In this paper, data from the typical meteorological year (TMY) for Belo Horizonte from the literature (SWERA, 2018) were used to determine a daily clearness index. A typical meteorological year (TMY) is an artificial year derived from multi-year weather datasets using mathematical methods (Yuan *et al.*, 2022). For each month, the average radiation over the period is determined, together with the average radiation in each month during the same period. The month with the average radiation most closely equal to the monthly average was chosen as the TMY data for that month. The selection of the real data (in a specific year) is better than the use of a simple average of the yearly data because it can underestimate the amount of variability.

3. MATERIALS AND METHODS

Forecasting techniques depend on the nature, quality, and time resolution of available data. Two fundamental approaches are generally used to develop those forecasts: casual and time series models. Causal models are used for developing cause-effect relationships between inputs and outputs. In contrast, time series models predict future values by regressing their previously observed values (Chaturvedi *et al.*, 2022). Finding the best forecasting model in time series analysis has always been a big challenge for decision-makers (Falatouri *et al.*, 2022). In this paper, three time series forecasting models were selected and used to predict the daily solar radiation data for the state of Minas Gerais: SARIMA, Holt-Winters, and LSTM.

SARIMA is an extension of the Autoregressive integrated moving average (ARIMA) model, one of the most used for forecasting linear time series. SARIMA is the seasonal ARIMA, which presents as a main advantage its ability to consider the seasonal behavior of stationary or non-stationary time series (Falatouri *et al.*, 2022). Holt-Winters is a traditional time-series approach, with characteristics of low dependence on historical data. It considers all previous values while giving weight to the most recent values. The model requires low data storage, is simple to use, and can be easily automated (Omar and Kawamukai, 2021). LSTM (Long Short Term Memory) is one of the most popular models used for short and long dependency in the data, keeping the error constant while solving the vanishing gradient problem by avoiding recent observations (Chacón *et al.*, 2023). The model can accurately classify, process, and forecast time-series data (Chaturvedi *et al.*, 2022).

The historical data of solar radiation used to feed the forecasting models were obtained from data collected by NASA satellites over 10 years, from 2004 to 2014 (Stackhouse, 2020).

The simulation was performed to estimate solar radiation data for 2015, and the results were compared to the real data obtained from NASA satellites for this year. For the city of Belo Horizonte, the results were also compared to the results from an analytical model.

The solar radiation in 90 cities evenly distributed in Minas Gerais and vicinities was predicted, using the same machine learning methods. These data were interpolated, and a solar radiation map was generated for the state of Minas Gerais. The interpolation data method used was the Inverse Distance Weighting, which uses known values to estimate an unknown value based on a given weight that makes the unknown value closer to its nearest neighbor (Shepard, 1968).

4. RESULTS AND DISCUSSION

This section is divided into two parts, as follows: The first section presents the key results obtained from this work. The second section discusses the findings with some considerations regarding the methodology and the models used.

4.1 Main results

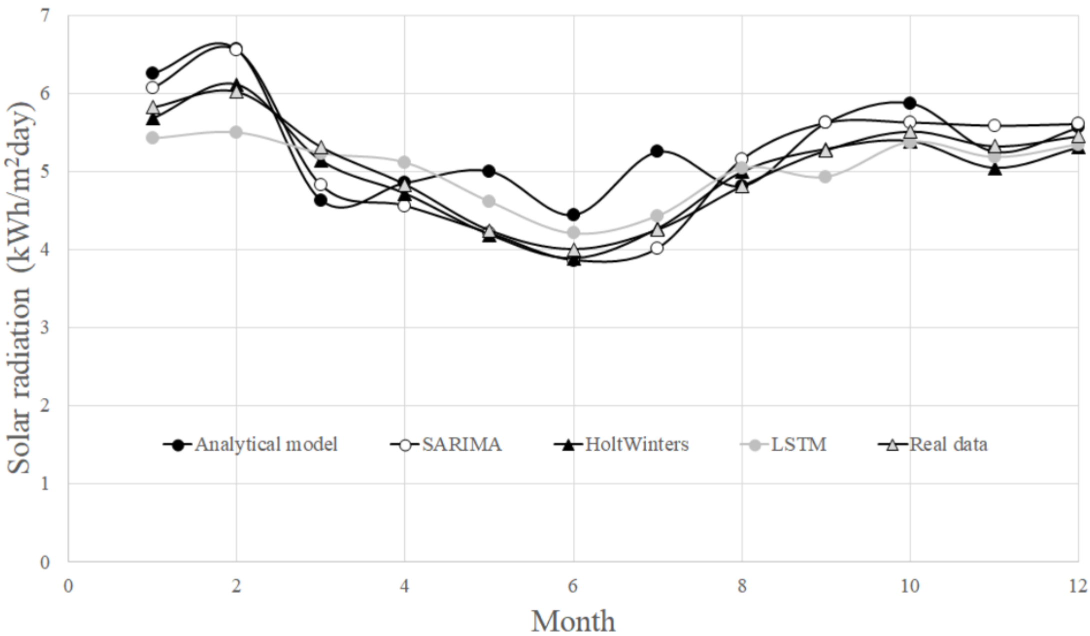


Figure 1: Predicted and real data for the solar radiation in Belo Horizonte.

The comparison of the average results obtained from the analytical model, from the three machine learning methods, and the real data from NASA satellites is shown in Figure 1.

Accordingly, both the analytical and numerical models performed well, with small differences compared to the real data. The analytical model presented the highest differences, with higher average and maximum percentual errors, compared to real data, as shown in Table 1. Among the machine learning methods, Holt-Winters obtained the best performance, with SARIMA and LSTM with similar performances.

Table 1: Average and maximum percentual error compared to real data for Belo Horizonte.

Method	**Average error**	**Maximum error**
Analytical model	8.2%	23.5%
SARIMA	5.1%	9.2%
Holt-Winters	2.4%	5.2%
LSTM	4.9%	8.8%

The same methodology was applied to predict the solar radiation for the state of Minas Gerais. The historical data used to feed the forecasting models comprises solar radiation data from 90 cities evenly distributed along the state collected from NASA satellites, from years 2004 to 2014. Table 2 summarizes the average values of solar radiation obtained using the same machine learning methods, as well as the average values of the solar radiation collected in the year 2015. The results are also presented in Figure 2.

Table 2: Average solar radiation for Minas Gerais (kWh/m2/day).

Method	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
SARIMA	6.10	6.21	5.25	4.86	4.48	4.01	4.24	5.29	5.56	5.78	5.67	5.70
Holt-Winters	5.70	6.00	5.36	4.95	4.42	4.10	4.41	5.17	5.39	5.54	5.41	5.59
LSTM	5.83	5.74	5.65	5.32	4.79	4.51	4.87	5.33	5.44	5.61	5.74	5.75
Real data	5.90	5.95	5.39	5.00	4.39	4.10	4.35	4.93	5.30	5.58	5.58	5.66

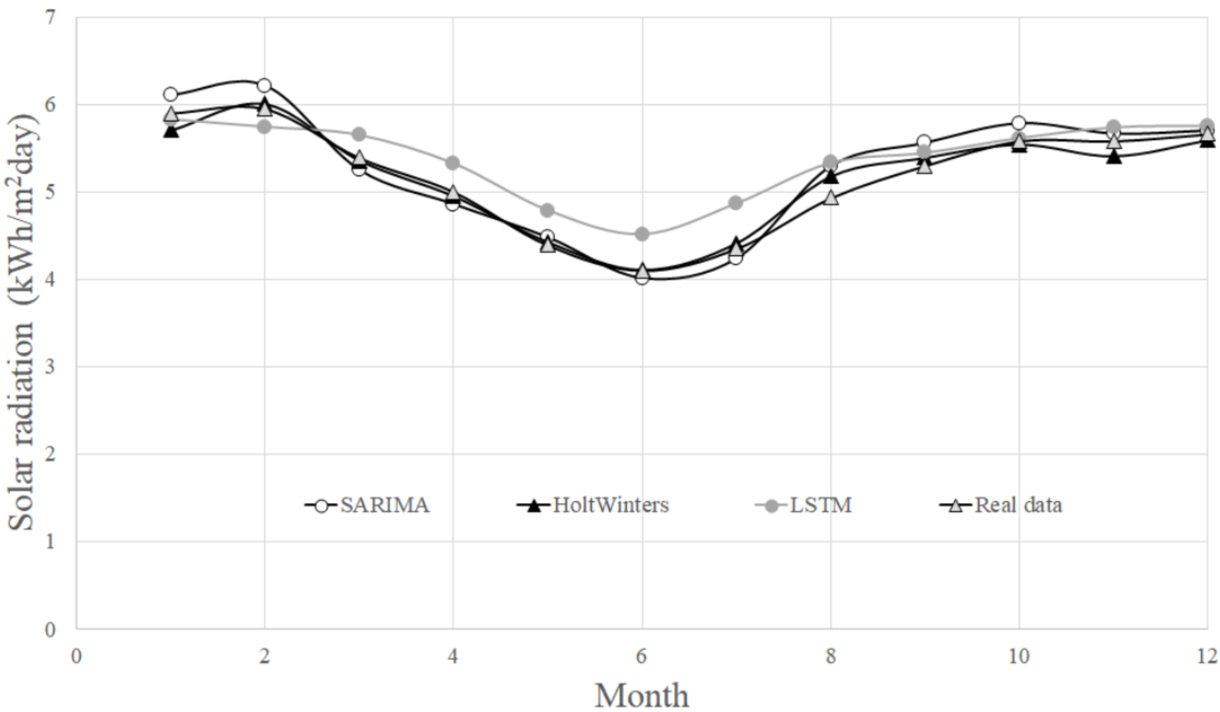


Figure 2: Predicted and real data for the solar radiation in Minas Gerais.

Table 3: Average and maximum percentual error compared to real data for Minas Gerais.

Method	**Average error**	**Maximum error**
SARIMA	3.2%	7.4%
Holt-Winters	1.6%	4.9%
LSTM	5.2%	11.9%

The average and maximum percentual errors of the machine learning methods, compared to real data for Minas

Gerais, are shown in Table 3. Holt-Winters presented the best performance, followed by SARIMA, and LSTM presented the highest differences.

The previously presented results are average data, used to define the machine learning with the best performance. Considering that Holt-Winters showed the lowest differences, this model was used to generate a solar radiation map for the state of Minas Gerais. We used monthly averaged data to create one map for each month, shown in Figure 3. The upper left map refers to January, and the lower right map refers to December.

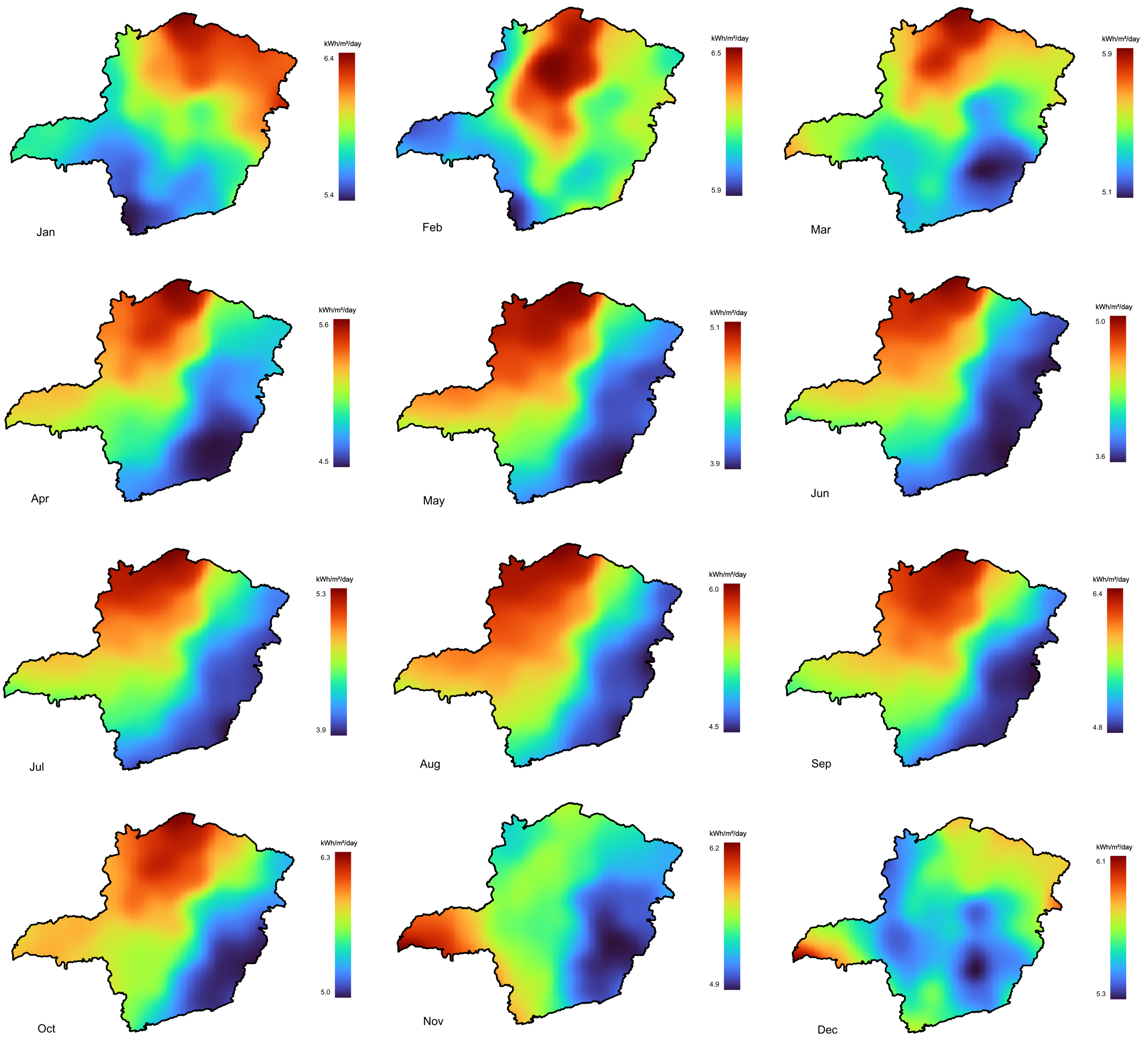
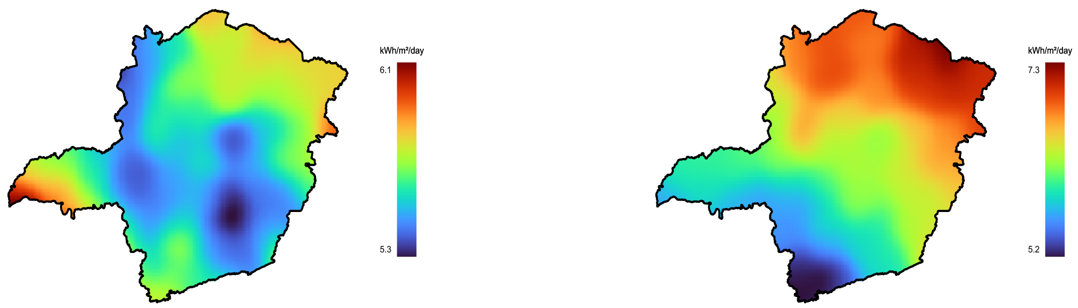


Figure 3: Monthly solar radiation maps for Minas Gerais

The comparison between the yearly averaged solar radiation data predicted by the Holt-Winters method and the real data obtained from NASA satellites for 2015 is shown in Figure 4. Figure 4a shows the predicted values and Figure 4b shows the real data. The numerical prediction model was able to successfully predict the real data.



(a) Holt-Winters (b) Real Data

Figure 4: Yearly averaged solar radiation map for Minas Gerais

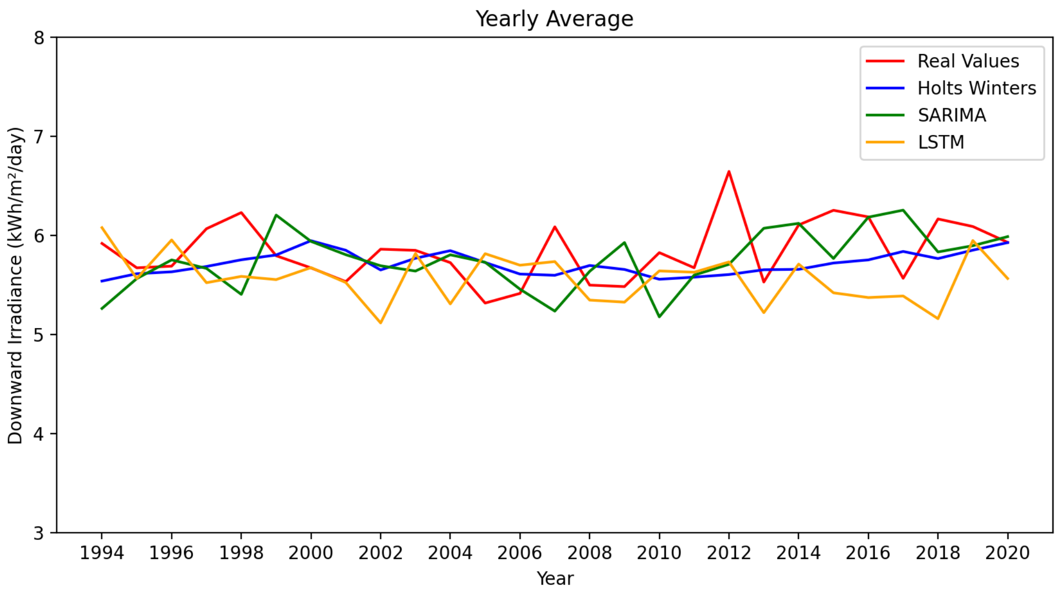


Figure 5: Yearly average variation of solar irradiance for the state of Minas Gerais.

Figure 5 shows the average yearly solar irradiance for the state of Minas Gerais from 1994 to 2020, to better visualize how each model adapted to the changes in the average solar behavior in a larger time frame.

4.2 Discussion

The solar radiation that reaches the surface of the earth varies due to variations in extraterrestrial radiation (caused by variations in the radiation emitted by the sun and by variations of the earth-sun distance) and to the random characteristics of the climatic conditions, caused by clouds, reflection, and absorption of the radiation, among other aspects. Therefore, in different years, it is expected to measure different values of solar radiation. For these reasons, it was not expected that the values obtained by the different models were the same.

The number of locations selected, 90, was suitable because it was possible to obtain different patterns of solar radiation. As expected, the distribution of solar radiation was different throughout the year. In general, the solar incidence was higher in the northern region of the state.

One key characteristic of solar behavior is that the absolute values of solar radiation that reach earth can vary significantly from year to year. In this regard, it is important to analyze how the machine learning models adapted to the variation of solar radiation over a more extensive period of time.

As shown in Figure 5, despite presenting the lowest error compared to the other models, the Holt-Winters model presents the slightest variation of yearly mean solar irradiance between the models; this can be attributed to the smoothing algorithm used in the model, which over a large time frame tends to centralize the solar irradiance predictions to values close to the average.

It is also worth pointing out that the LSTM model presented lower yearly averages compared to the actual values of solar radiation.

Although the models did not accompany the actual average of solar radiation, all three models were capable of consistently predicting future values comparable to the actual solar behavior of the studied time frame. Given that the three models were trained using a single variable - Solar irradiance- the inclusion of other variables, such as clearness index, relative humidity, and atmospheric temperature, can significantly improve the models' prediction capabilities, resulting in lower error.

5. CONCLUSIONS

In this paper, three machine learning methods (SARIMA, Holt-Winters, and LSTM) were used to predict solar radiation in the city of Belo Horizonte, Brazil. The results were compared to results from an analytical model and to real data obtained from NASA satellites. The results were obtained for the year 2015. The analytical model presented the highest differences related to experimental data. Among the numerical prediction models, Holt-Winters showed the best performance.

Using satellite data from 90 cities in the state of Minas Gerais and vicinities, the machine learning methods were also used to predict the solar radiation in these cities. When comparing the average results with real data for 2015, Holt-Winters presented the best performance. The solar radiation data were interpolated to generate a solar radiation map for the state of Minas Gerais, for each month of the year. The yearly averaged solar radiation map was compared to real data, and it was concluded that the three models were able to successfully predict solar radiation, with the Holt-Winters model presenting the lowest error.

The methodology can be applied to forecast solar radiation for future years, helping the design of solar equipment. The methodology can be applied to forecast other variables, such as ambient temperature and wind speed if data are available. It is worth mentioning that these and other variables' data are freely available on the NASA website.

6. Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. The authors are also thankful to Puc Minas, CNPq, and FAPEMIG.

7. REFERENCES

Budiyanto, M.A. and Lubis, M.H., 2020. "Physical reviews of solar radiation models for estimating global solar radiation in indonesia". *Energy Reports*, Vol. 6, pp. 1206–1211.

Chacón, H., Koppisetti, V., Hardage, D., Choo, K.K.R. and Rad, P., 2023. "Forecasting call center arrivals using temporal memory networks and gradient boosting algorithm". *Expert Systems with Applications*, Vol. 224, p. 119983.

Chaturvedi, S., Rajasekar, E., Natarajan, S. and McCullen, N., 2022. "A comparative assessment of sarima, lstm rnn and fb prophet models to forecast total and peak monthly energy demand for india". *Energy Policy*, Vol. 168, p. 113097.

Duffie, J.A. and Beckman, W.A., 2013. *Solar engineering of thermal processes*. John Wiley & Sons.

Falatouri, T., Darbanian, F., Brandtner, P. and Udokwu, C., 2022. "Predictive analytics for demand forecasting–a comparison of sarima and lstm in retail scm". *Procedia Computer Science*, Vol. 200, pp. 993–1003.

Formolli, M., Kleiven, T. and Lobaccaro, G., 2023. "Assessing solar energy accessibility at high latitudes: A systematic review of urban spatial domains, metrics, and parameters". *Renewable and Sustainable Energy Reviews*, Vol. 177, p. 113231.

Kosovic, I.N., Mastelic, T. and Ivankovic, D., 2020. "Using artificial intelligence on environmental data from internet of things for estimating solar radiation: Comprehensive analysis". *Journal of Cleaner Production*, Vol. 266, p. 121489.

Kumar, C.M.S., Singh, S., Gupta, M.K., Nimdeo, Y.M., Raushan, R., Deorankar, A.V., Kumar, T.A., Rout, P.K., Chanotiya, C., Pakhale, V.D. *et al.*, 2023. "Solar energy: A promising renewable source for meeting energy demand in indian agriculture applications". *Sustainable Energy Technologies and Assessments*, Vol. 55, p. 102905.

Omar, M.S. and Kawamukai, H., 2021. "Prediction of ndvi using the holt-winters model in high and low vegetation regions: A case study of east africa". *Scientific African*, Vol. 14, p. e01020.

Pandey, A.K., Kumar, R. and Samykano, M., 2022. "Solar energy: direct and indirect methods to harvest usable energy". In *Dye-Sensitized Solar Cells*, Elsevier, pp. 1–24.

REN21, 2021. "Facts from the renewables 2021 global status report: Trends in brazil." URL `https://www.ren21.net/wp-content/uploads/2019/05/REN21`$_{G}SR2021_{F}actsheet_{B}razil_{E}N.pdf$.

Shepard, D., 1968. "A two-dimensional interpolation function for irregularly-spaced data". In *Proceedings of the 1968 23rd ACM national conference*. pp. 517–524.

Stackhouse, P., 2020. “Nasa powerl data access viewer”. *URL: power. larc. nasa. gov*.

Staff, I., 2019. “Co2 emissions from fuel combustion: 2019 edition”. *International Energy Agency*.

SWERA, 2018. URL `https://openei.org/wiki/Solar`$_{a}nd_{W}ind_{E}nergy_{R}esource_{A}ssessment_{(}SWERA)$.

Yolcan, O.O., 2023. “World energy outlook and state of renewable energy: 10-year evaluation”. *Innovation and Green Development*, Vol. 2, No. 4, p. 100070.

Yuan, J., Huang, P. and Chai, J., 2022. “Development of a calibrated typical meteorological year weather file in system design of zero-energy building for performance improvements”. *Energy*, Vol. 259, p. 125031.

8. RESPONSIBILITY NOTICE

The authors are solely responsible for the printed material included in this paper.