Düzce University
Journal of Science & Technology

Research Paper

Monthly Electricity Generatıon Forecast in Solar Power Plants with LSTM[1]

Ömer ÇETİN [a,*], Ali Hakan IŞIK [b]

[a] *Bilgisayar Mühendisliği Bölümü, Mühendislik Mimarlık Fakültesi, Mehmet Akif Ersoy Üniversitesi, Burdur, TÜRKİYE*
[b] *Bilgisayar Mühendisliği Bölümü, Mühendislik Mimarlık Fakültesi, Mehmet Akif Ersoy Üniversitesi, Burdur, TÜRKİYE*
** Corresponding author's e-mail address: eng.omercetin@gmail.com*
DOI: 10.29130/dubited.1015251

ABSTRACT

Today, with the intensive use of electrical devices, the need for electricity has increased. Fossil fuels are generally used to meet this need. However, considering the damage caused by fossil fuels to the environment, governments make various incentives for renewable energy sources. The incentives of countries for solar power plants are quite large. Recently, there are many investors who want to build solar power plants. The sunshine duration of our country is quite high. And the fact that the climatic conditions are efficient for the generation of electricity attracts many investors. However, the installation of these power plants is quite costly. It is possible to predict the amortization periods of these costs with the ever-developing artificial intelligence technology. In this study, the energy data to be produced in the future is estimated by using real solar power plant data with machine learning algorithms. Data, take from solar power plants owned by Humartaş Energy company. In the study, predictions and analyses were made using the LSTM (Long Short-Term Memory) method, which is one of the artificial neural networks. The error rate of the study between 1% and 15%. It is foreseen that studies will also be implemented with other renewable energy sources such as wind, geothermal, hydraulic energy data in the coming stages.

Keywords: *LSTM, Artificial intelligence, Solar power plants, Energy forecast*

LSTM ile Güneş Enerjisi Santrallerinde Aylık Elektrik Üretim Tahmini

ÖZ

Günümüzde elektrikli cihazların yoğun kullanımı ile elektriğe olan ihtiyaç artmıştır. Bu ihtiyacı karşılamak için genellikle fosil yakıtlar kullanılmaktadır. Ancak fosil yakıtların çevreye verdiği zararı göz önünde bulundurarak hükümetler yenilenebilir enerji kaynakları için çeşitli teşvikler yapmaktadır. Ülkelerin güneş enerjisi santrallerine yönelik teşvikleri oldukça fazladır. Son zamanlarda güneş enerjisi santrali kurmak isteyen birçok yatırımcı var. Ülkemizin güneşlenme süresi oldukça yüksektir. İklim koşullarının elektrik üretimi için verimli olması da birçok yatırımcıyı cezbetmektedir. Ancak bu santrallerin kurulumu oldukça maliyetlidir. Sürekli gelişen yapay zekâ teknolojisi ile bu maliyetlerin amortisman sürelerini tahmin etmek mümkün. Bu çalışmada, makine öğrenmesi algoritmaları ile gerçek güneş enerjisi santrali verileri kullanılarak gelecekte üretilecek enerji verileri tahmin edilmektedir. Veriler, Humartaş Enerji firmasına ait güneş enerjisi santrallerinden alınmıştır. Çalışmada yapay sinir ağlarından biri olan LSTM (Uzun Kısa Süreli Bellek) yöntemi kullanılarak tahmin ve

[1]ICAIAME 2021 konferansında sunulmuştur.

Received: 27/10/2021, Revised: 18/11/2021, Accepted: 22/11/2021

analizler yapılmıştır. Çalışmanın hata oranı %1 ile %15 arasındadır. Önümüzdeki aşamalarda rüzgâr, jeotermal, hidrolik enerji gibi diğer yenilenebilir enerji kaynakları ile de çalışmaların yapılması öngörülmektedir.

Anahtar Kelimeler*: *LSTM, Yapay Zekâ, Güneş Enerji Santralleri, Enerji Tahmini

I. INTRODUCTION

In Turkey, which continues to grow and develop, there is a continuous increase in energy consumption day by day due to increasing industrial activities, developing technology and ever-increasing population. In response to this increasing consumption, solutions are tried to be produced with fossil fuels. However, due to the high carbon emissions of fossil fuels, the damage to the environment is considerably higher. The International Energy Agency report predicts that there will be a 20% increase in energy-related CO_2 emissions in 2035 [1]. The vast majority of carbon emissions are caused by fossil fuels. The Paris Agreement was signed as part of the United Nations Framework Convention on Climate Change in 2015. With the agreement that came into force in 2016, Turkey started to reduce fossil fuel consumption by increasing investment in renewable energy sources. The process was accelerated by giving significant incentives to investments. Along with these investments, solar power plants have also benefited greatly from the incentives. The average sunshine duration in our country, which is measured as 7 hours and 50 minutes, shows that the amortization period of solar power plants is shorter compared to other countries [2]. Distribution of installed power by resources in Turkey as of the end of September 2019; 31.4 percent hydraulic energy, 28.6 percent natural gas, 22.4 percent coal, 8.1 percent wind, 6.2 percent solar, 1.6 percent geothermal and 1.7 percent is in the form of other sources. Solar power plants constitute 6,2% of this power, and this rate is increasing day by day.

Artificial intelligence, which emerged in 1956 and progressed with the development of technology, appears in many parts of our daily lives. Artificial intelligence is the imitation of human features such as learning, decision making and prediction by machines. In short, it is imitate of the human brain by machines. Personal assistants used in phones are a simple example of artificial intelligence application [3].

We make our work easier with electrical energy in all areas of our lives. Artificial intelligence helps us to use time efficiently. A study has been carried out to combine these two fields, which are popular today and will continue to be popular in the future. To increase the accuracy of the software, machine learning algorithms were written based on the data taken from the energy company. Monthly average electricity production estimates of the power plants to be installed have been made using the weather conditions, sunshine duration, power plant location, solar panel power, inverter power, albedo effects of the environment of solar power plants. To make predictions, the LSTM method from artificial neural networks was used. Having the ability to store data in memory compared to other artificial neural networks increases the predictive power. Using this advantage, it is aimed to estimate with a high accuracy rate. In addition, as a result of the study, investors with a solar power plant will be able to find out the amount of electricity generation a month in advance. Investors who want to install a solar power plant due to the large exchange rate differences and increasing material and labor costs are wondering about the depreciation periods. With this study, the depreciation periods of the solar power plants to be installed can be easily found out. It is targeted to develop a software for data collection operations in the later steps of the study. In this way, it is purposed to make an application with high predictive power by continuously training the data.

II. SYSTEMS USED IN SOLAR POWER PLANTS AND FACTORS AFFECTING THE EFFICIENCY OF POWER PLANTS

Photovoltaic batteries consist of modules with a certain number of serial and parallel connections [4]. Although more efficient photovoltaic cells are used in space technologies, there are many photovoltaic cells today.

There are many cell types in photovoltaic modules [5-6]:

(i) Crystal and multi-crystal silicon solar cells
(ii) Thin-film solar cells (a-Si, cadmium telluride and CIS)
(iii) Nanotechnology-based solar cells (tandem, super tandem and dec solar cells, etc.)

The most commonly used types are crystalline silicon, cadmium telluride, and CIS/CIGS cells [7].
Solar power plants are formed as a result of combining photovoltaic modules with suitable angles against the sun. Many factors affect the working conditions of power plants. Parameters such as the nominal characteristic values of the components that make up the plant, the system configuration, the location of the plant, the reflection effect of structures around the plant area cause various losses.

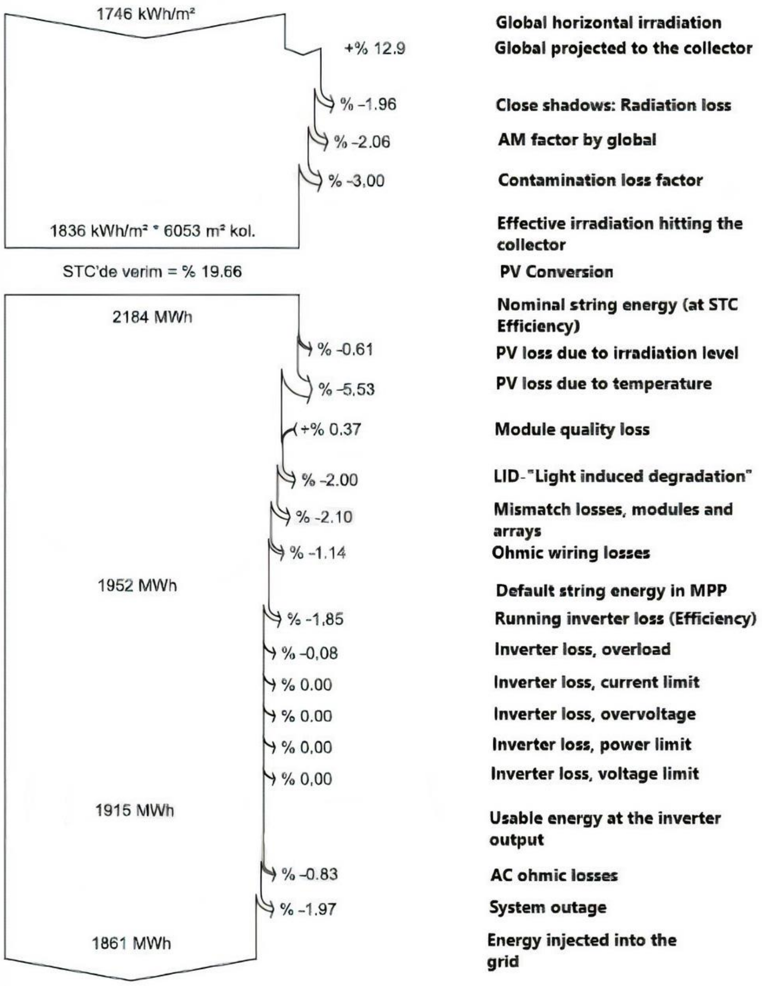

Figure 1. *Loss rates in solar power plant*

Companies that build solar power plants generally use the PVSyst program. With this program, the performance rates of the plant to be installed are known. Figure 1 shows an example of a PVSyst report. In this report, which factors generally cause energy loss in solar power plants, are seen.

A. RADIATION LOSSES

Depending on the angle of the photovoltaic (PV) module, the effect of general photons can increase or decrease. In Turkey, this angle varies between 25 and 30 degrees depending on the location.

Depending on the location of the solar power plant, the sunshine duration changes. The angle of arrival of the sun's rays is important, as well as the duration of sunbathing.

If there is a deviation from the standard solar spectrum in AM 1.5, in the case of instantaneous radiation, losses occur in the spectral response of the PV modules in the system [7]. PV module selection should be made by considering these parameters.

Spectral response sizes of CIS (Copper Indium Selenide) and c-Si (Crystal Silicon) technologies are wider than other technologies. In this way, their spectral absorption is higher. Solar panels made of CdTe (Cadmium Telluride) and a-Si materials have limited spectral response areas ranging from 350 to 800 nm. Modules made of a-Si material produce more electrical energy at large solar angles and high solar radiation [7]. Figure 2 shows this much more clearly.

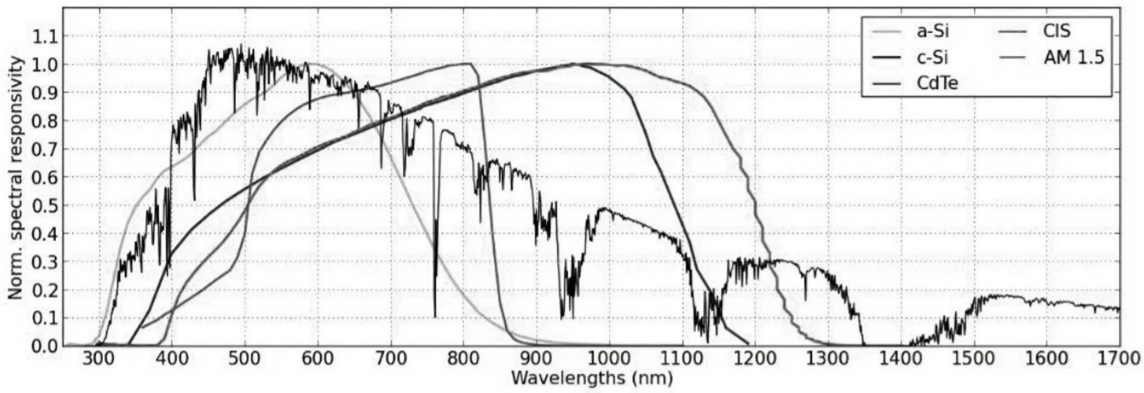

Figure 2. *Normalized spectral responses of PV technology*

The spectra of the sky with solar radiation observed over one year differ from the ideal spectrum (AM 1,5) of the clear sky. Depending on this, sunshine duration is an important factor for solar power plants.

In daily life, the fact that the sun is rarely blocked by clouds is an important factor for electricity generation. When the clouds come in front of the sun, radiation loss occurs. As a result, it causes a loss of efficiency in the PV modules.

Solar power plant investors and engineers designing the system, choose PV modules depending on the location of the plant by in consideration at the PV module characteristics. The important parameters in the catalogs include power temperature coefficient (γPmpp), Standard Test Conditions- module output power at STC (Watt Peak- Wp), efficiency at STC (ηSTC), and module parameters with lower irradiance values than STC [8].

B. WEATHER

Another parameter that affects solar power plants is the weather. PV modules work ratio inversely with temperature. With the heating of the cells, the panel efficiency decreases. When we look at the solar map in our country, the power plant efficiency is higher in cities that are generally cold but have high sunshine duration. The temperature of a PV cell under the conditions of 1 m/s wind, 800 W/m2 radiation, and 20 °C ambient temperature is called the Nominal Operating Cell Temperature (NOCT). The air temperature at which a solar panel is most efficient is 25 ° C, as shown in figure 3. Building a solar power plant by considering this temperature parameter will increase the efficiency of the plant.

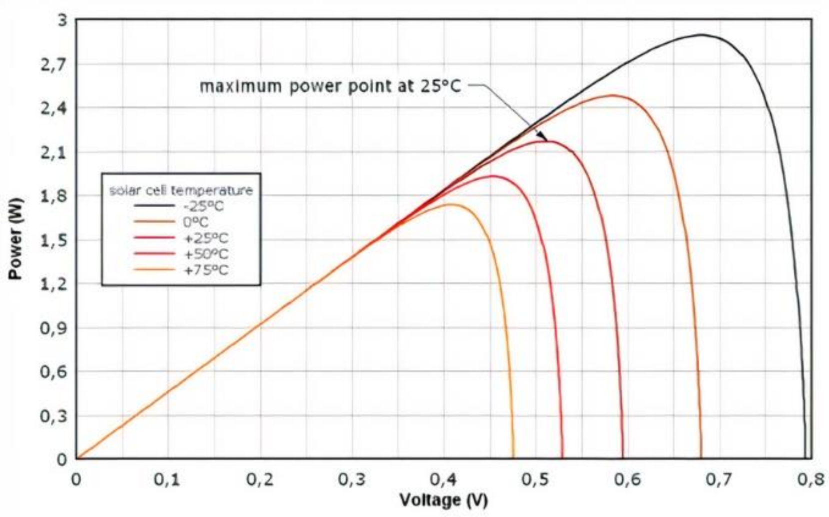

Figure 3. *Temperature dependence of PV panels [9]*

C. INVERTER EFFECT

Inverters are electrical power converters that convert direct current (DC) to alternating current (AC). During this cycle, there is a loss of efficiency due to the electrical difference. DC/AC conversion losses in the inverter; depend on the type of semiconductor, power layer topology used in the PV inverter, and the usage characteristics of the capacitors such as conduction, magnetic elements, and switching [10].

In general, PV inverter efficiency decreases by 0.3%-1% for every 150V DC input voltage amplitude. However, at high DC input voltages, a decrease of up to 5% is observed in the power consumption of the control unit, low radiation, and switching losses [11]. In today's technology, European transformerless grid-connected inverters have a maximum efficiency of 98% at nominal DC input voltages.

D. ALBEDO EFFECT

Some of the sun rays falling on the PV module are reflected back without absorbing by the panels. Therefore, reflection losses occur.

The radiation falling on a solid body usually performs the following three optical motions:

• Reflection: The return of the radiation on the surface of the object.
• Conduction: Penetration of radiation into the object.
• Absorption: Penetration and capture of radiation into the object; energy is converted into a different form.

In PV modules, some of the radiant flux is absorbed and converted into electrical energy. The quality of the reflective material varies according to the smallness of the losses in transmission and the intensity of the absorbed radiation. Crystalline silicon absorbs more light than amorphous silicon. The modules are produced in different layers, preventing reflection and maximizing radiation absorption. PV modules have a tempered structure designed to minimize surface coating reflection and maximize the absorption of glass. PV cells are covered with an anti-reflective coating material that prevents the reflection of light.

The intensity of the reflected solar radiation changes according to the refractive index of the module and the angle between the sun and the PV module. In normal radiation, the PV modules reflect 4% of the incident light [12].

III. FORECASTING SPP GENERATION DATA WITH LSTM METHOD

Classical deep learning networks work independently of input and output data. But RNN (Recursive Neural Network) can influence the output using the input data. Just as people can form a complete sentence based on their previous experience in writing a sentence, the RNN artificial neural network can also make predictions and make the words come in a logical sequence. RNN tries to keep the results of previous steps in its memory. Theoretically, RNNs are expected to give good results in long sequences, but as a result of practical experience, it has been seen that they cannot achieve this. Networks such as LSTM and GRU (Gated Recurrent Units) have been designed to overcome these problems.

Long Short Term Memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Like other neural networks, it works with feedback logic, not feed-forward logic. Not just one data point, such as image processing, but all data series, such as speech and video, can be processed. An ordinary LSTM unit consists of 4 layers. These are a cell, an entrance door, an exit door and a forget door. The cell remembers values at random time intervals, and these three gates regulate the flow of information into and out of the cell.

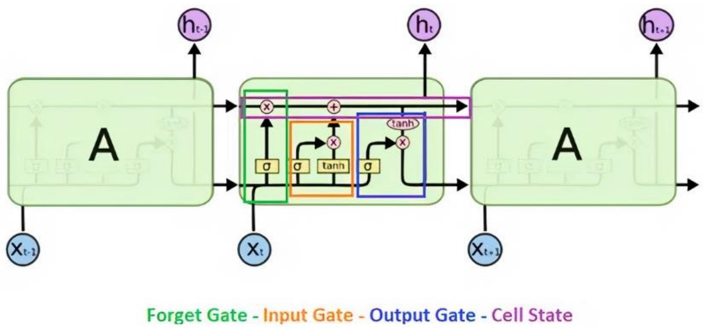

Figure 4. *LSTM structure [13]*

The Cell State cell carries meaningful information across cells to make predictions. This cell can be explained as a communication line and memory of the network. This solves the short-term memory problem and historical data can be moved on the network. Gates can determine the data that the Cell State cell carries. Thanks to the gates, necessary or unnecessary data can be distinguished. The data coming to the gates are compressed between 0 and 1 with the sigmoid activation function. As a result of the activation, the data that is 0 is forgotten. The data that is 1 continues to progress with the Cell State cell.

The forget gate decides which information will be forgotten and which will continue. The information (ht) from the previous cell and the current information (xt) are inserted into the sigmoid activation function and the decision is made as a result. While the information with 0 is forgotten, the information with 1 continues to be carried with the Cell State.

The entrance gate updates the Cell State. It puts current information and previous information into the sigmoid activation function. As a result, it decides whether to update or not. 0 information is unimportant. It uses information that is 1. At the same time, the tanh activation function, which compresses the data to the range of -1 and 1, is used for the regulation process. Then, the sigmoid and tanh function outputs are multiplied and it is decided which information to update.

The output gate determines the next cell's input (ht+1). It is also used to make predictions. First, current information and previous information are passed from the sigmoid function. Then the existing information on the Cell State is passed through the tanh function. Finally, the two results are multiplied and it is decided which information will be the input (ht+1) cell for the next cell. When the gate operations are completed, the Cell State that will go to the next cell and the Hidden State(ht) information, which is defined as the cell's login information, is decided.

In the study, new electricity generation data were estimated in a time series based on previous data using the LSTM method. The data were taken from an energy company and selected from the factors affecting energy loss in solar power plants. Energy data history (month-year), weather, sunbathing time, panel technology, panel power, inverter power, power plant age, albedo values of the city where the power plant is located, and the region where the power plant is located were taken into account and monthly electricity production was estimated in the future. Similar studies have been conducted using the LSTM method in previous studies [14]. But since this study was not conducted specifically for the location of Turkey, it becomes difficult to use this study in Turkey. For this reason, wanted to conduct a study specific to Turkey. Depending on the location of Turkey, different parameters were considered and energy data were estimated using a more original and only LSTM method.

In order for the data to be trained faster and the process to be completed faster, coded in Google Colabs. Python was used as the programming language.

Data from power plants located in nine cities were used in the study. The monthly data includes how much electricity the power plant produced in 2020. Nearly a thousand data were used for training.

IV. RESULTS AND ANALYSIS

Having a lot of training data in the monthly electricity generation forecast study at solar power plants will increase the accuracy rate and more efficient results will be achieved. In order for there to be no overfitting learning, test and train rates must be adjusted very well in learning. Otherwise, there is a high probability of encountering incorrect forecasts.

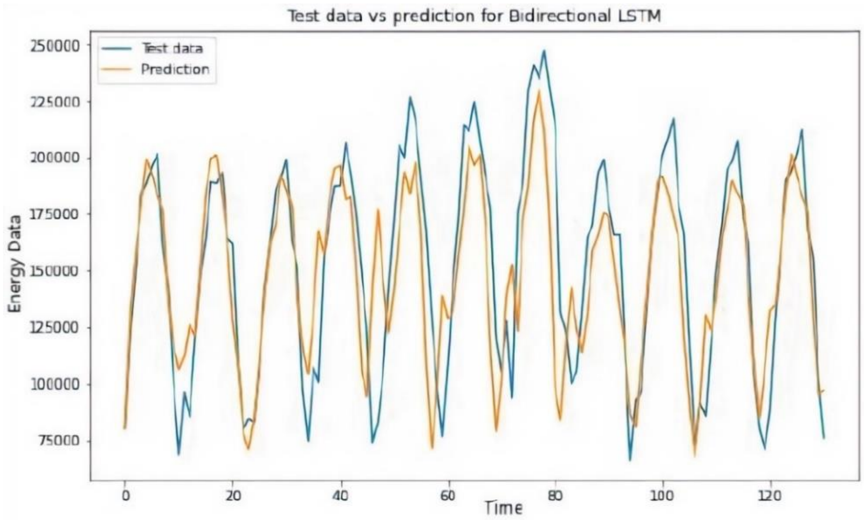

Figure 5. *Estimation with the LSTM method*

Figure 5 shows the estimation graph made by the LSTM method. The actual data and forecasts of a 120-Month city are displayed. Although in some months the forecast data is very close to the actual data, it is clearly visible that there is a deterioration after the 40th month. Especially in summer, when the energy data is increasing, it is observed that there are differences between the forecast data and the actual data. In Figure 6, the loss rates are seen more clearly.

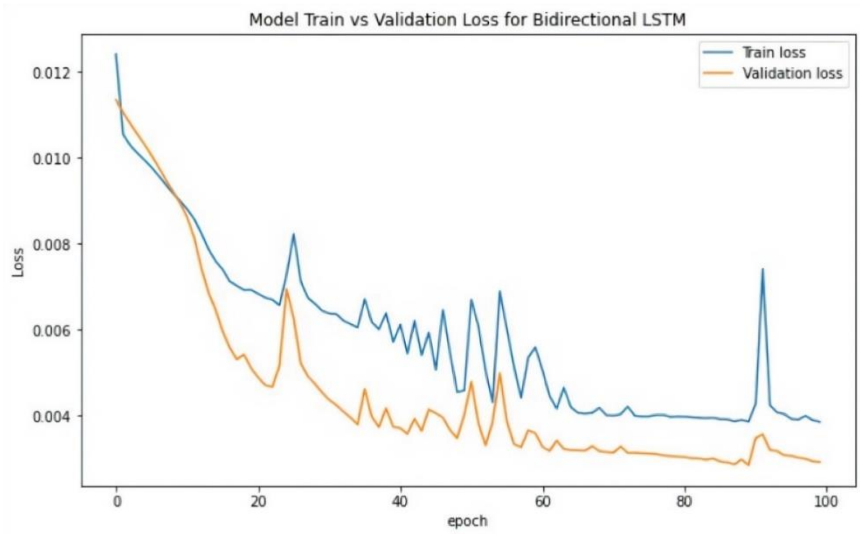

Figure 6. *LSTM method loss rate*

Missing data makes it difficult to analyze data, process data, and may cause efficiency to decrease [15]. In order to avoid overfitting that may occur due to low data, data duplicate was performed using the Data Imputation Method.

```
prediction_bilstm.tolist()

[[80490.34375],
 [132247.875],
 [159890.0],
```

Figure 7. *Energy data obtained as a result of estimation*

2020.01	91337.31
2020.02	95706.96
2020.03	139865.95

Figure 8. *Real time and field energy data*

$$ER = \frac{(Actual\ Data - Forecast\ Data)}{Forecast\ Data} x100 \quad (1)$$

ER=Error Rate

$$\text{January Error Rate} = \frac{(91337.31 - 80490.34)}{800490.34} x100 \quad (2)$$

The error rate for January is approximately 13% according to the 2nd Equation calculation. Considering this value for the month of May, it appears as 1%. Values vary widely. This is thought to be due to the scarcity of data.

V. CONCLUSION

In the presented study, LSTM management from artificial neural networks was used. It has been tried to estimate the amount of electricity that the power plant will produce in the next month with the help of past production data of solar power plants from renewable energy sources. As a result of the study, it was seen that great success was achieved very close to reality in the estimation of energy data using the LSTM neural network method. With the study, it is possible to predict how much electricity will be produced in the coming periods by investors who want to invest in solar power plants or who have a solar power plant. This method can be applied to most renewable energy plants. Thanks to this, the production volumes at the power plants can be estimated a month in advance. If an example is given, it is concluded that an energy production forecast can be made depending on the wind speed, weather, wind blade material at wind power plants or the annual precipitation rate at hydroelectric power plants. According to the results of the study, the error rate was found to be between 1% and 15%. It has been concluded that this result remains much lower compared to other artificial intelligence studies.

VI. REFERENCES

[1] Word Energy Outlook 2013, (2013). *World Energy Outlook 2013* [Online]. Available: www.iea.org/reports/world-energy-outlook-2013.

[2] T.C. Enerji ve Tabii Kaynaklar Bakanlığı, (2020). *Güneş* [Çevrimiçi]. Erişim: www.enerji.gov.tr/bilgi-merkezi-enerji-gunes.

[3] F. Özen, (2019, Haziran). *Yenilenebilir Enerjide Yapay Zeka Uygulamaları* [Çevrimiçi]. Erişim: www.researchgate.net/publication/333816757_YENILENEBILIR_ENERJIDE_YAPAY_ZEKA_UYGULAMALARI.

[4] T. Boyekin, I. Kıyak, "Rooftop solar power plant based electric vehicle charging station," in *6. European Conference On Renewable Energy Systems (ECRES)*, 2018, pp. 959-966.

[5] H. Hemza, C. Abdeslam, M. P. Rachid and D. Barakel, "Tracing current-voltage curve of solar panel based on LabVIEW Arduino interfacing," *Bilişim Teknolojileri Dergisi*, vol. 8, no. 3, pp. 117–123, 2015.

[6] V. Ciocia, A. Boicea, A. Dematteis, P. D. Leo, F. Giordano and F. Spertino, “PV system integration in buildings: an energy and economic case study,” in *2017 10th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, 2017, pp. 786-790.

[7] P. Grunow, A. Preiss, S. Koch and S. Krauter “Yield and Spectral Effects of A-Si Modules,” in *Proceedings of the 24th European Photovoltaic Solar Energy Conference*, 2009, pp. 2846-2829.

[8] B. Herteleer, J. Cappelle and J. Driesen, “Quantifying low-light behaviour of photovoltaic modules by identifying their irradiance-dependent efficiency from data sheets,” in *European Photovoltaic Solar Energy Conference*, pp. 1-8, 2012.

[9] Solar PV Word Expo, (2021, August). *Solar Cells* [Online], Available: https://www.pvresources.com/en/solarcells/solarcells.php.

[10] E. Koutroulis, F. Blaabjerg, “Methods for the Optimal Design of Grid-Connected PV Inverters, International Journal Of Renewable Energy Research,” *IJRER*, vol. 1, no. 2, pp. 54-64, 2011.

[11] B. Giesler, “String vs. Central inverters: Dimension of the inverter,” in *Photon’s 1st PV Inverter Conference*, 2010, pp. 1-23.

[12] Oğuzhan, (2019). *Fotovoltaik Sistemlerin Verimliligin Etkilen Faktorler* [Çevrimiçi]. Erişim: www.enerjisistemlerimuhendisligi.com/fotovoltaik-sistemlerin-verimliligin-etkilen-faktorler.html.

[13] O. Akköse, (2020, 22 Aralık). *Uzun-Kısa Vadeli Bellek (LSTM*) [Çevrimiçi]. Erişim: https://medium.com/deep-learning-turkiye/uzun-k%C4%B1sa-vadeli-bellek-lstm-b018c07174a3.

[14] A. Gensler, J. Henze, B. Sick and N. Raabe, “Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM Neural Networks,” in *2016 IEEE İnternational Conference on Systems, Man, and Cybernetics (SMC)*, 2016, pp. 2858-2865.

[15] S. M. Susanti, E. Sulistianingsih, “K Nearest Neighbor dalam Imputasi Missing Data,” *Bimaster*, vol. 7, no. 1, pp. 9-14, 2018.