Solar Power Prediction using LTC Models

Anunay Gupta¹, Anindya Gupta², Apoorv Bansal³ and Madan Mohan Tripathi⁴

¹²³⁴Electrical Engineering, DTU, New Delhi, India

*Correspondence: Madan Mohan Tripathi; Email: mnntripathi@gmail.com

ABSTRACT - Renewable energy production has been increasing at a tremendous rate in the past decades. This increase in production has led to various benefits such as low cost of energy production and making energy production independent of fossil fuels. However, in order to fully reap the benefits of renewable energy and produce energy in an optimum manner, it is essential that we forecast energy production. Historically deep learning-based techniques have been successful in accurately forecasting solar energy production. In this paper we develop an ensemble model that utilizes ordinary differential based neural networks (Liquid Time constant Networks and Recurrent Neural networks) to forecast solar power production 24 hours ahead. Our ensemble is able to achieve superior result with MAPE of 5.70% and an MAE of 1.07 MW.

Keywords: Deep learning, Ensemble, Neural Ordinary differential equations, Solar energy forecasting.

1. INTRODUCTION

There has been an increase in the demand for energy which has led to an increase in the emission of harmful agents in the atmosphere in large amounts such as CO₂. Due to this, the need and usage of renewable sources of energy have increased in order to preserve non-renewable sources of energy and also bring down the pollution level in our environment [1]. These sources replenish at an equal or a faster rate at which they are being consumed. Apart from the fact that renewable sources of energy are in abundance and can be replenished easily, they help to save non-renewable fuels like coal, oil, etc. They are also less of a threat to the environment [2] which solves many issues including air pollution, global warming, etc. They also help in cutting down the cost of the production of energy and have the opportunity to boost the efficiency of its generation. Due to the aforementioned reasons, the countries are continuously promoting renewable energy and investing heavily in it and expect it to overtake conventional energy sources in power generation and transportation which are the major areas of energy consumption. Some of the prominent sources are wind, solar, geothermal, etc.

The most common and widely used form of renewable energy is solar energy. It is the heat and energy coming from the sun and it is used in various places such as solar power which in turn generates electricity. Another application of solar energy is solar thermal energy which is used in solar architecture and solar water heating. Many countries are finding various ways to incorporate solar energy into usage, reducing the use of depletable energy. India alone increases its annual Cumulative capacity by around 33% from year 2019 to 2020 and is expected to meet its 40% of energy requirements via non-renewable sources of energy by the end of 2022. Solar energy not only is clean energy and helps save the atmosphere from harmful agents but also cuts down the cost of production. PV grid systems are very cheap to maintain [3].

Earth receives a huge amount of solar energy and a big part of it is absorbed by the cloud and otherwise reflected back to space. Earth receives so much solar energy that it overtakes the amount of energy that is used by the world by a huge margin. Though the energy is huge, there is an unequal distribution of it according to the geography of the location. The locations which are nearer to the equator receive the greatest amount of energy and as we go farther from the equator, the intensity reduces and its potential in various locations differs as well. Although, if the countries farther away from the equator use photovoltaics which helps focus towards the Sun by following it, its potential automatically increases. Other factors that affect the use of solar energy are time variation and cloud cover. Time variation affects solar energy because there is minimal radiation and hence it is most effective during the daytime. Cloud cover is also an important factor because if the clouds cover the sun, the intensity of incoming sunlight will be reduced and hence the generation of solar power will be affected.

Another notable factor is land availability. It is needed to ensure places are available for the installation of solar panels. People in households prefer to install them on their roofs as they have now realized that they can directly store energy from their homes. Solar power is generated by converting sunlight energy into electricity. We can achieve this by using photovoltaics or using concentrated solar power. For the purpose of small and medium-sized applications, photovoltaics was used in the old days. Mirrors and lenses are now used by concentrated solar power stations for focusing huge amounts of sunlight into a single beam.
Solar energy depends on various factors such as wind speed, cloud cover, temperature, solar radiation intensity, etc. These factors sometimes lead to an inconsistent supply of energy to the grid [4]. This can lead to grid failure and cause harm to other equipment. To prevent this, a backup generator is used so that there is no margin between the energy requirement and the energy produced. Hence it is vital to forecast the PV energy produced beforehand which can help plant setup planners and grid operators to have an idea about the photovoltaic energy present and the energy required. This helps in a cost-effective setup of the grid and efficient management of the energy produced [5].

Another basis of distinction can be the method used which includes physical, statistical, artificial intelligence, or a hybrid method. The statistical method is based on the use of historical data to forecast. An association is developed between the variable to be predicted and the input data [15]. This is not proven to be an accurate method as AI methods and hybrid methods provide much higher accuracy. A physical model uses mathematical equations to forecast. They depend on the characteristic of the photovoltaic, atmospheric variable important to the forecast, geographical location, etc. [4]. These models are usually complex and slow due to the time consumed in the computation of the equations. A hybrid method is comprised of various individual models which themselves can be used to forecast solar power [19] and has proven to show better results than the ANN model. Very vivid research is being carried out in the field of forecasting solar power using various machine learning models. Due to the quick learning ability of such models and AI models, they are used widely.

In authors of [16] used 7 various individual models to generate three different ensemble models for day ahead forecasting. The results show that such a hybrid model performs and shows more accurate results than any individual model alone. In [17] authors used forecasting methods based on ANN showed that it is very accurate and efficient when compared to the original data. This model also withstands changes in weather conditions and improves the results significantly. Another model to predict daily predict photovoltaic energy is using the BP algorithm on a neural network-based model [18].

There have been numerous efforts in forecasting renewable energy using various machine learning algorithms such as Long Short-Term Memory (LSTM), Support Vector Machine (SVM), and K-nearest Neighbour (KNN). The authors of [7] uses a hybrid Gradient boosted regression tree (GBRT) model which comprises various similar-sized regression trees. It was used to forecast solar power generation for a future time period of 1-6 hours in multiple regions. This model doesn’t find extensive use because of its limitations of updating procedures as new data is collected. A Least Absolute Shrinkage and Selection Operator (LASSO) model was used in [8] for forecasting solar power. LASSO reduced the weightage of less important factors. This model was compared to SVM and TLE (time-series method) and LASSO proved to be forecasting better and with fewer training datasets. LSTM was merged with the LASSO forecasting model and it showed better results for the short-term forecasting rather than long-term forecasting.

In [9] authors developed a model using the AutoEncoder (AE) and LSTM neural network. The model showed quite accurate results in forecasting the solar power when the encoding part of AE was used to feed the LSTM the most effective features as input. The authors of [10] used a slightly different approach and merged the model in [9] with a forecasting model using

2. LITERATURE REVIEW
Notable progress has been made in the field of forecasting solar energy. The most important reason for this is to reduce the uncertainties in the production and supply of energy to the grid. Accurate forecasting can help the operators to keep in check with the happenings of the extreme event in order to protect the grid from it. This also helps to calculate the extra energy sources required to maintain a consistent flow of energy and also reduces the cost of the system and increases efficiency [6]. Figure 1 shows us the time horizons of solar energy forecast.

So solar energy forecasting can be differentiated on the basis of the time period for which forecasting is to be made. This includes very short term, short term, medium-term, and long-term forecasting [14].

Figure 1: Time horizons of solar energy forecast
augmented long short-term memory (A-LSTM). Results were quite fascinating on time-series datasets.

In [11], many ML models used in forecasting various renewable energy systems were studied. As a result of comparison among the 6 models namely, Neural Network (NN), Multiple Linear regression (MLR), K-NN, SVN, and 2 other models, it was found that the Neural Network Ensemble’s models showed more precise outputs.

There are two types of ensembles forecasting methods as per the authors in [12] namely Cooperative and Competitive Ensemble forecasting. For forecasting the day ahead solar power, various ensemble techniques are mentioned in [13]: -
1. Linear
2. Normal Distribution
3. Normal Distribution with Additional features

After several comparisons, various authors discovered that hybrid models work better than any individual forecasting model e.g., SVM, K-NN, etc. under any weather circumstances. They show more accurate and promising results showing that diversification of model is important to achieve better results.

Building upon this observation we introduce a novel ensemble model which utilises neural ordinary differential equations based liquid time constant networks and traditional recurrent neural networks to predict day ahead solar power with great efficiency. The major contributions of this paper are summarized below:
1. Proposed a novel ensemble model to predict day ahead solar power.
2. Showed the efficacy of Liquid Time constant networks in predicting solar power.
3. Compared the performance of LTC’s and LTC based ensembles with traditional recurrent neural networks in predicting day ahead solar power.

The rest of the paper is organised in three sections namely proposed architecture, Experimental results and Conclusion and future works. In proposed architecture we dive deep into the details about the function of proposed model, in experimental results section we analyse the results obtained by various models as well explain in detail about the data and metrics used in the experiment. The conclusion and future works section presents our findings and the works that can be done in future to further improve forecast models.

### 3. PROPOSED MODELS

Traditional sequential networks like RNN’s are robust deep learning models that effectively learn the general trend of a time series. However, when it comes to renewable energy, the time series of such data has a lot of variations and is composed of many peaks and troughs due to changes in weather variables like solar irradiance, cloudiness, wind speed, etc. That traditional forecasters generally fail to capture such drastic variation. To capture these variations, we decided to introduce liquid time constant networks (LTC). Liquid time constant networks are touted for their better expressivity than traditional sequential networks and hence can capture more complex trends present in time series. Our proposed architecture as seen in Figure 2 consists of three main constituents, namely: 1) LTC-Based encoder, 2) RNN-Based encoder 3) FC-Based stacked Decoder. Before discussing about the architecture, we will look into the dataset that we used for our work.

#### 3.1 LTC Based Encoder

Liquid time constant networks harness the power of neural ordinary differential equations to gain high expressivity. To appreciate the innovation in LTC it is essential to understand the functioning of neural ordinary differential equations.

Neural ordinary equations attempt to learn the function that governs the transformation of inputs to the desired output. While traditional neural networks can be thought of as group of finite number of non-linear layers in which each layer transforms the input discretely and drastically. In contrast Neural ordinary equations can be thought as having infinite layered networks in which the input undergoes continuous transformation and all such transformations are controlled by one function.

To better understand what differentiates a neural ordinary differential equation from traditional neural networks, let’s look at eq(1) which shows a single resnet layer. In a Resnet layer the input of the layer is transformed by multiplying the inputs by weights and adding input to create a skip connection. Thus each layer of a resnet learns from the difference between its input and output. Neural differential equation extrapolate the process of learning the change between input and output seen in networks like Resnets. As seen in eq(2) if we are to replace 1 with constant Δ and we then limit Δ to near 0 as shown in eq(3), we reach eq(4), the principle equation that governs neural ordinary differential equation. Through eq(4), the continuity of neural ordinary differential equation becomes more evident. In neural ordinary differential equations we try to accurately estimate dh/dt which controls how the transformation of input to output will occur. Once dh/dt is finalized, we estimate the dependent variable using ODE solvers.

\[
\begin{align*}
&\Rightarrow h(t+1) = ActivationFunction(W(t)h(t) + b(t)) + h(t) \quad (1) \\
&\Rightarrow \frac{h(t+\Delta) - h(t)}{\Delta} = f(h(t), \theta(t)) \quad (2) \\
&\Rightarrow \lim_{\Delta \to 0} \frac{h(t+\Delta) - h(t)}{\Delta} = f(h(t), \theta(t), t) \quad (3) \\
&\Rightarrow \frac{dh(t)}{dt} = f(h(t), \theta(t), t) \quad (4)
\end{align*}
\]

Liquid Time Constant networks are more expressive and accurate neural ordinary equations. As shown in eq(5) in LTC's,
the state is computed using an ode which has two major parameters: 1) damping constant (\( \tau \)) and a specialized input \( S(t) \) which is basically the product of the function of neural network and difference of hidden states and a bias factor as shown in eq(6). \( S(t) \) plays a critical role in defining LTC’s behaviour since due to \( S(t) \), the coefficient of state \( x(t) \) becomes dependent on the function of the neural network \( F \), which is in turn parametrized by the time step. Thus, LTC can adapt to different dynamic systems at each time step.

In our proposed architecture, as shown in fig (2) each LTC cell receives a hidden state (a 50-feature tensor) from the previous LTC cell and input at that timestep. LTC cells then harness neural ode's to append the information gained from input at current timestep in the hidden state and pass the hidden state to the next LTC cell. This way LTC cell outputs will have information gained from all timesteps. The output from the last time step is fed into fully connected layer. The outputs of this FC layer are then fed into fully connected stacked decoder.

\[
\frac{dx(t)}{dt} = \frac{-x(t)}{\tau} + S(t) \quad (5)
\]

\[
S(t) = f(x(t), I(t), t, \theta)(1 - x(t)) \quad (6)
\]

\[
\frac{dx(t)}{dt} = \left[1 + f(x(t), I(t), t, \theta)\right] x(t) + f(x(t), I(t), t, \theta) \quad (7)
\]

### 3.2 RNN Based Encoder

The second encoder in our proposed architecture is RNN based encoders. The RNN based encoder lends robustness to our model. LTC's though highly expressive is also highly variant and is susceptible to overfitting and noisy data. A RNN based encoder placed in parallel to LTC encoder solves such problems. In our proposed architecture the RNN-based encoder is composed of multiple RNN-cell. Each RNN cell receives an input corresponding to its timestep and a hidden state from the previous RNN cell. The RNN cell then takes the weighted sum of input and the hidden state received to create the next hidden state. The output from last timestep is the culmination of inputs at all timesteps. This output is fed into a fully connected (FC) dense layer. The output of this FC layer is then fed into fully connected stacked decoder to get merged with the output of the LTC-Based encoder.

\[
h(t + 1) = f_H(W_{HH} x(t) + W_{HH} h(t)) \quad (8)
\]

\[
y(t + 1) = f_O(W_{HO} h(t + 1)) \quad (9)
\]

### 3.3 Fused Stack Decoder

The FC stacked decoder serves the purpose of merging the insights gained from both LTC encoder and RNN encoder and accurately predicting the target variable. The Fully stacked decoder consists of 3 dense layers. Firstly, the inputs received from LTC encoder and RNN encoder are concatenated to form FC layers' input. The concatenated vector is then passed to the above mentioned FC layers which have the following dimensions \((53, 26); (26, 13); (13, 1)\) after going through each layer ReLU activation is applied to introduce non-linearity. The output of the last FC layer is the target variable which for the study is day ahead prediction of generated power.

### 4. EXPERIMENTAL RESULT

#### 4.1 Dataset Description

The dataset we used for predicting solar power is taken from Pagwada solar plant. This plant is located in a small city located in the Tumkur district, Karnataka. The solar power production data was taken over a period of 2 years starting from 30th March 2020 to 23rd April 2021 at an interval of 10 minutes. We averaged the power production data into an interval of 1 hour using. This corresponded to 8904 steps of 1hr each. In addition to solar power produced we also utilised weather data such solar radiation, temperature and hour of the day. The weather parameters were obtained from NASA power access data viewer for the district of Pagwada. The selection of the above-mentioned feature was done using analysis by co-relation we chose features which had high correlation with solar power produced and had very less correlation amongst themselves.

#### 4.2 Evaluation Metrics

We evaluated the performance of the various models using these evaluation parameters; Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). MAPE is the only parameter in which percentage is used while in the others, absolute values are used.

MAPE is the average of the percentage error in forecasting values of solar power. If for the actual power \( A_p \), the model predicts the power as \( P_p \), then MAPE is given in equation 10 as:

\[
MAPE = \frac{100}{n} \sum_{p=1}^{n} \frac{|A_p - P_p|}{A_p} \quad (10)
\]

MSE is defined as the average of the squared error in predicting valued of solar power. If for the actual power \( A_p \), the model predicts the power as \( P_p \), then MSE is given in equation 11 as:

\[
MSE = \frac{1}{n} \sum_{p=1}^{n} (A_p - P_p)^2 \quad (11)
\]

MAE is similar as MSE but the difference between them is that instead of taking squared error, MAE takes absolute error. If for the actual power \( A_p \), the model predicts the power as \( P_p \), then MAE is given in equation 12 as:

\[
MAE = \frac{1}{n} \sum_{p=1}^{n} |A_p - P_p| \quad (12)
\]

RMSE is obtained when root of the MSE of the data is taken. If for the actual power \( A_p \), the model predicts the power as \( P_p \), then MSE is given in equation 13 as:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{p=1}^{n} (A_p - P_p)^2} \quad (13)
\]

#### 4.3 Results

In this section we analyse and compare the predictions made by various models both qualitatively by means of line graphs of
predictions and quantitively by means of various metrics such as MAPE, MSE, MAE and RMSE.

Table 1 summarises the results obtained by models in day ahead solar power forecast of Pagwada Solar plant.

<table>
<thead>
<tr>
<th>Models</th>
<th>MAPE</th>
<th>MSE</th>
<th>MAE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTC</td>
<td>7.010 ± 6.92</td>
<td>2.975 ± 4.886</td>
<td>1.321 ± 1.10</td>
<td>1.72</td>
</tr>
<tr>
<td>LSTM</td>
<td>7.015 ± 8.20</td>
<td>3.247 ± 5.25</td>
<td>1.389 ± 1.14</td>
<td>1.80</td>
</tr>
<tr>
<td>RNN</td>
<td>6.13 ± 8.40</td>
<td>2.37 ± 5.35</td>
<td>1.12 ± 1.05</td>
<td>1.53</td>
</tr>
<tr>
<td>LTC + LSTM</td>
<td>6.11 ± 7.30</td>
<td>2.39 ± 4.85</td>
<td>1.14 ± 1.04</td>
<td>1.54</td>
</tr>
<tr>
<td>LTC + RNN</td>
<td>5.70 ± 6.50</td>
<td>2.19 ± 5.06</td>
<td>1.07 ± 1.02</td>
<td>1.48</td>
</tr>
</tbody>
</table>

On analysing the table, it becomes clear that the ensemble of LTC and RNN prove to be the most effective and achieves an MAPE of 5.70%, MSE of 2.19, MAE of 1.07 and RMSE of 1.48. The reason behind such effective results is the fact that LTC and RNN cater to different problems as explained in proposed architecture section, LTC owing to the use of neural ordinary differential equation is very effective in adapting to variations. This can also be seen in the results obtained by LTC. Though LTC achieve 7.010% mean MAPE it achieved best standard deviation of 6.92 amongst non-ensemble model. On the other hand, RNN achieves one of the best MAPE of 6.13% amongst non-ensemble models, however it has the highest standard deviation which support our hypothesis explained in proposed architecture section that RNN effectively captures the general trend. One of the interesting results that we come across while analysing this dataset is that LSTM fails to perform significantly better than other networks. LSTM achieves a MAPE of 7.015 which is slightly worse than LTC, but in contrast to the case of LTC, LSTM also does not achieve a good standard deviation. The findings of the metrics are reiterated by the qualitative metrics. Figure 3- Figure 7 depicts the predication graph of all models on a 100-hour window of test data. It can be observed that prediction made by RNN though comes very close to actual values in non-peak regions. LTC on the other hand are able deal with the peak region well when compared to RNN. The ensemble of LTC and RNN perform the best and produce the best visual predictions.
5. CONCLUSION

In this study, we propose an ensemble model to predict day ahead solar power. The proposed ensemble model is composed of two major components LTC and RNN's which provide different benefits to this ensemble model. While RNN is great at capturing general trend of the series, it lacks the ability to capture the frequent variations present in solar power data, this is where LTC comes into play since LTC is based on neural ordinary differential equations. It has lot of expressive power leading to LTC capturing the frequent variation and deviations in solar power produced. The proposed ensemble model is able to efficiently predict solar power achieving an MAPE of 5.7%. Another avenue that we explore while conducting this study is to probe the efficiency of LTC in the domain of solar power predictions LTC prove to be powerful networks and achieve slightly better results than LSTM's and much better results in terms of standard deviation of metrics. In the future in order to further improve solar power predictions better features such as clear sky index can be integrated in the data. On the network front neural ODE’s which form the basis liquid time constant networks can be modified to make LTC more stable and less prone to noisy data.

REFERENCES


© 2022 by Anunay Gupta, Anindya Gupta, Apoorv Bansal and Madan Mohan Tripathi. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).