Long-term Solar Radiation Forecasting using a Deep Learning Approach-GRUs

Muhammad Aslam
Kim Hyung Seung
Seung Jae Lee
Dept. of Electrical Engineering
Myongji University
Yongin, South Korea
aslammju@gmail.com

Jae-Myeong Lee
Sugwon Hong
Eui Hyang Lee*
Dept. of Computer Engineering
Myongji University
Yongin, South Korea
ljm9317kr@gmail.com
*Next-Generation Power Technology Center (NPTC)

***Abstract*— A long-term solar generation forecasting is an important issue in a microgrid design. Solar generation forecasting mainly depends on solar radiation forecasting. In this paper, A Deep Learning approach using Gated Recurrent Units (GRUs) is proposed for forecasting of a year-ahead hourly and daily solar radiation. The proposed GRU model is compared with the state of the art methods like Long Short Term Memory (LSTM), Recurrent Neural Network (RNN), Support Vector Regression (SVR), Feed Forward Neural Network (FFNN) models and numerical model. Its effectiveness for long-term solar radiation forecasting over other methods is verified.**

Keywords*— *Deep Learning, Microgrid, Renewable energy, Solar Radiation forecasting, Gated Recurrent Unit, Long Short Term Memory

I. INTRODUCTION

Renewable energies are playing vital role in energy sector primarily due to the ability to combat global warming, economical and diversifying energy mix ensuring energy security and sustainability. Solar generation is the most common types of renewable energy that has grown rapidly over the past decade and it is expected to grow even faster in the future [1], [2]. Even though comprehensively increasing in the size and the installed capacity, the uncertainty and variability of renewable generation poses big challenges. Also, to help grid's operation with planning, maintenance and operation, energy economics quantities like the electricity price should be forecasted [3]. Following benefits can be obtained by an accurate renewable generation forecast: firstly help to carry out planning and maintenance, secondly minimize penalties and charges due to imbalance of generated power and thirdly provide a good knowledge of future energy market trading [4]. Specially, long-term solar power generation forecasting is essential from engineering and market point of view. It is necessary to estimate renewable generation capacity, Energy Storage System (ESS) capacity, total demand, simulation capacities, and micro-grid market participation. For these purposes, long-term generation and demand needed to be forecasted. Numerous parameters effect the solar generation forecasting, but solar irradiation is the key component for solar generation. Therefore, long-term solar irradiance is needed to be forecasted correctly.

Stochastic differential equation model is used to derive probabilistic forecast of solar irradiance during a day at a given location [5]. Similar work has been done to estimate long-term solar radiation forecasting [6]. Online forecasting of PV generation is approached suiting for many applications and used to predict for thirty-six hours [7]. Artificial Intelligence (AI) specifically Deep Learning (DL) for PV and wind power generation is comprehensively reviewed concluding its benefit [8-9]. LSTM is used for year-ahead hourly and daily long-term solar radiation forecasting [10]. In this paper, GRU is used for year-ahead hourly and daily solar radiation forecasting. A comparison of GRU and LSTM along with comparison with RNN, FFNN and SVR are also made. The paper is organized as follows: Section II explains the Deep Learning architectures used for time series forecasting, Section III explains proposed models, data and experimental results. Section IV gives conclusion.

II. DEEP LEARNING ARCHITECTURES SUITABLE FOR TIME SERIES

In this section, FFNN, RNN, LSTM and GRU architectures are explained. Section II-A will describe the Feed Forward Neural Network (FFNN) while section II-B will describe the Recurrent Neural Network (RNN) and section II-C will explain a comparison of Recurrent Neural Network (RNN) extensions i.e. LSTM and GRU suitable for time series forecasting.

A. Feed Forward Neural Network (FFNN)

Figure 1 shows a typical FFNN. It is simplest form of a deep neural network where the input is fed to the hidden layer in forward direction until output is calculated using activation functions in each hidden layer node and initialized weights and biases. Later, weights are adjusted using backpropagation algorithm using the loss functions to get optimal values.

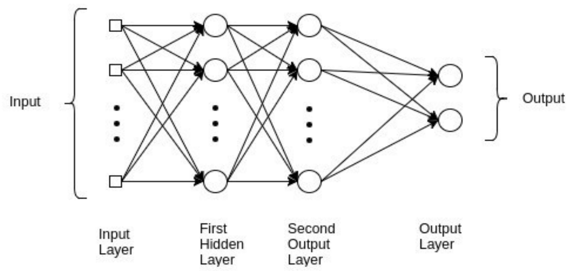

Figure 1. An example of FFNN

B. *Recurrent Neural Network (RNN)*

1) Definition: An exemplary RNN network is shown in Figure 2. An RNN is an extension of a conventional FFNN, which is able to handle a variable-length sequence input. RNN is thought to be useful for sequence data as it handles the variable-length sequence by having a recurrent hidden state whose activation at each time is dependent on that of the previous time-step.

RNN consists of one hidden layer which maintains hidden state and a fully-connected layer to get output from its hidden states. And hidden state h_t is defined as:

$$h_0 = 0$$

$$h_t = \tanh(Ux_t + Wh_{t-1}) \quad (1)$$

Where, x_t is input on hidden layer and h_{t-1} is previous hidden state from the previous time-step, and U is weight matrix for input, and W is weight matrix for previous hidden state.

After hidden state is obtained, RNN output o_t is directly calculated from current hidden state in a fully-connected layer.

$$o_t = \sigma(Vh_t) \quad (2)$$

Where, σ is sigmoid activation function.

2) Problem to RNN and Solutions: Unfortunately, it has been observed [11] that it is difficult to train RNNs to capture long-term dependencies because the gradients tend to either most of the time vanish or rarely tend to explode. To solve these issues LSTM unit was introduced which is followed by Gated Recurrent Units (GRUs) introduced more recently [12]. RNNs applied using these units have shown to solve the issues of gradient by capturing long-term dependencies.

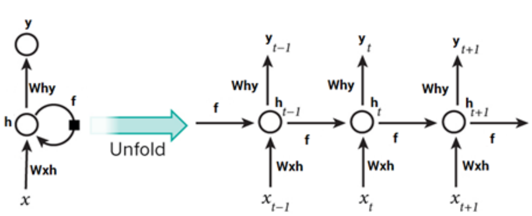

Figure 2. Recurrent Neural Network (RNN)

C. *LSTM vs GRU*

Unlike to the recurrent unit which simply computes a weighted sum of the input signal and applies a nonlinear function, each LSTM unit maintains a memory at a time. An LSTM unit consists of three gates i.e. input, output and forget with internal memory. An LSTM unit is able to decide whether to keep the existing memory via the introduced gates or forget. Intuitively, if the LSTM unit detects an important feature from an input sequence at early stage, it easily carries this information over a long distance, hence, capturing potential long-distance dependencies.

The input gate of LSTM decides how much current information needs to be passed.

$$i_t = \sigma(x_t U^i + h_{t-1} W^i) \quad (3)$$

The forget gate decides about the information need to be passed from previous state and is defined as

$$f_t = \sigma(x_t U^f + h_{t-1} W^f) \quad (4)$$

The output gate defines the internal state information need to be passed.

$$o_t = \sigma(x_t U^o + h_{t-1} W^o) \quad (5)$$

The internal memory also known as cell state C_t is updated in two steps. At first, a candidate of cell state $\tilde{C}_t$ is calculated by the following formula:

$$\tilde{C}_t = \tanh(x_t U^g + h_{t-1} W^g) \quad (6)$$

After $\tilde{C}_t$ is calculated, cell state C_t is updated from $\tilde{C}_t$.

$$C_t = \sigma(f_t * C_{t-1} + i_t * \tilde{C}_t) \quad (7)$$

Where, $*$ is operation of elementwise multiplication.

The cell state Ct is used to get hidden state of next time-step by the following formula:

$$h_t = \tanh(C_t) * o_t \quad (8)$$

Similar to the LSTM unit, the GRU has gating units i.e. reset gate and update gate that modulate the flow of information inside the unit, however, without having a separate memory cells.

The reset gate of GRU is defined as:

$$r_t = \sigma(x_t U^r + h_{t-1} W^r) \quad (9)$$

The update gate output z_t is calculated by the following formula:

$$z_t = \sigma(x_t U^z + h_{t-1} W^z) \quad (10)$$

The hidden state of GRU h_t is updated in two steps. Firstly temporary output of hidden state from reset gate is calculated as $\tilde{h}_t$. After that, h_t is calculated with update gate output z_t.

$$\tilde{h}_t = \tanh(x_t U^h + (r_t * h_{t-1}) W^h)$$
$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (11)$$

The GRU does not have any mechanism to control the degree to which its state is exposed, but exposes the whole state each time. Therefore, GRUs just expose the full hidden content without any control. It is computationally more efficient with less complexity. It is observed to perform faster and better than LSTM on certain data.

Comparison of LSTM vs GRU units is shown in Figure 3.

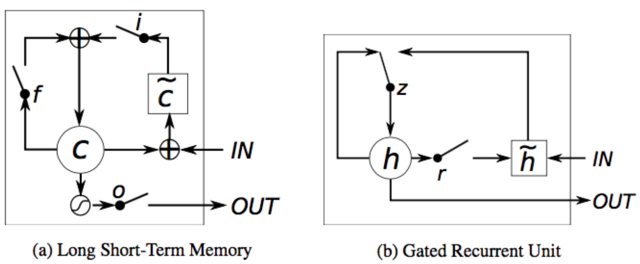

Figure 3. Comparison of LSTM vs GRU units

III. Proposed Model, Data and Results

A. The case study, data and models

As discussed earlier, in order to estimate renewable energy capacity, ESS capacity, ESS operation simulation in design engineering and market benefits, it is necessary to predict the amount of generation and load before system installation. Total power generation of one year is required in order to estimate capacity for designing. Therefore, it is necessary to predict the hourly and daily power generation for one year as well as the total power generation for one year in order to estimate the optimum capacity considering the operation of PV and ESS in systems like microgrid. Consequently, it is necessary to predict solar radiation accurately as it has a huge influence on the prediction of PV power generation.

The data are real-time solar radiation data of Seoul Korea from 2001 to 2017 obtain from Korea Meteorological Administration (KMA). In this paper, GRU is implemented for hourly and daily solar radiation forecasting for year-ahead. The model is compared with the state of the art models like SVR, RNN, FFNN and LSTM. All DL models consists of an input layer, three hidden layers with 32 nodes each and an output layer. The models was implemented in Python with Anaconda Jupyter notebook with Keras and TensorFlow at back-end. The error criteria used in this paper is Root Mean Square Error (RMSE).

B. Results

Table I shows the comparison of RMSE of different models. As can be seen from the Table I, GRU produces minimum RMSE proving its efficiency.

TABLE I. RMSE of Solar Irradiance for Different Models

Year	SVR	FFNN	RNN	LSTM	GRU
2016	4.8204	4.9423	5.3857	4.8175	**4.7936**
2015	5.3084	5.2552	5.5243	5.2000	**5.0880**

Table II shows performance of LSTM and GRU models. The performance is measured as training time in a system with AMD Ryzen Threadripper 2950X and 64GB RAM, and only CPU is used for model fitting and prediction. The measurements are mean values taken from 10 runs for accurate result. The result shows that the model GRU has higher performance compared to LSTM.

TABLE II. Performance of LSTM vs GRU

Year	LSTM (seconds)	GRU (seconds)
2016	76.38	**62.09**
2015	65.10	**56.93**

TABLE III compares total irradiation for one year ahead. From comparison it can be observed that the proposed method predicted data is similar to the actual data.

TABLE III. Total Yearly Radiation

Year	Actual (MJ/m²)	SVR (MJ/m²)	FFNN (MJ/m²)	RNN (MJ/m²)	LSTM (MJ/m²)	GRU (MJ/m²)
2016	4520.84	4552.05	4290.54	4389.38	4455.20	**4529.74**
2015	4647.27	4758.74	4276.76	4745.04	4501.91	**4610.41**

Figure 4 is a diagram showing absolute value of prediction errors for each model in the perspective of Total Yearly Radiation. The lower value indicates it is more similar to the actual data.

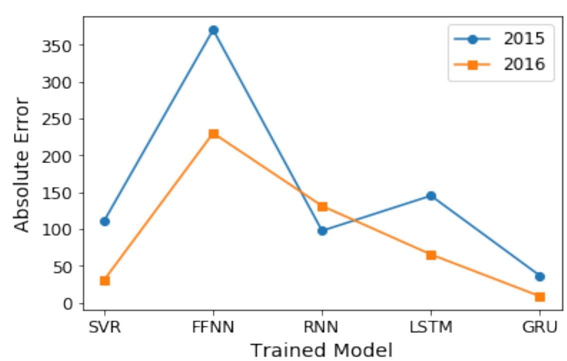

Figure 4. Absolute errors of trained models in Total Yearly Radiation

Figure 5 shows comparison of actual data vs vanilla RNN and its extensions (i.e. LSTM, GRU) trained with the data until year 2016, for monthly time-steps of year 2017.

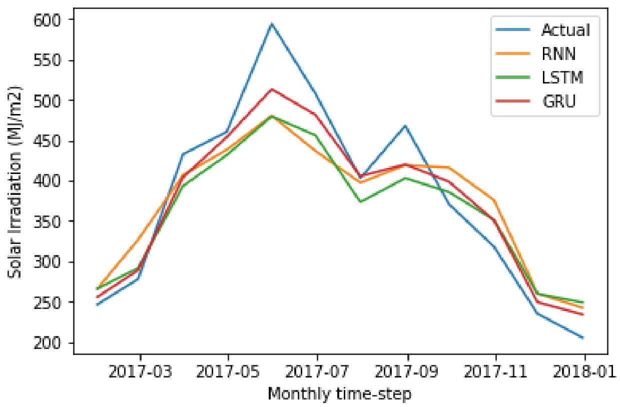

Figure 5. Comparison of RNN and its extensions in monthly time steps

In this test, we presented that GRU performs best on our data set compared to the other models as Table I shows, and also presented that GRU is faster than LSTM in our test system as shown in Table II. GRU performed around 14%~23% faster than LSTM.

For deeper observation, analyzed the prediction output from models in terms of Total Yearly Radiation and Monthly Radiation in Table III and Figure 4~5. Table III and Figure 4 showed that GRU predicted best in the yearly sum, and Figure 5 showed data similarity between models over months in the year 2017. Figure 5 proved prediction from GRU is most similar to the actual data.

IV. Conclusion

DL Long-term solar power generation forecasting is essential from engineering and market point of view. It is necessary to estimate renewable generation capacity, Energy Storage System (ESS) capacity, total demand, simulation capacities, and micro-grid market participation. For these purposes, long-term generation and demand need be forecasted. Solar power generation forecasting mainly depends on the amount of solar radiation. Previously, numerical models and LSTM recurrent network were used for the long-term i.e. year-ahead solar radiation forecasting and among them LSTM outperformed numerical methods. Recently many studies on time series report effectiveness of GRU over LSTM in terms of speed and accuracy and in this paper the GRU-based Deep Learning approach is adopted for a year-ahead daily and hourly solar radiation forecasting to see its effectiveness over LSTM and numerical models. The GRU model is also compared with FFNN, RNN and SVR, proving its efficiency. Future work could be to use extension of LSTMs and GRUs combined with auto-encoders for further improvement of accuracy both in short term and long term PV generation forecasting.

Acknowledgment

This work was supported by “Human Resources Program in Energy Technology” of the Korea Institute of Energy Technology Evaluation and Planning (KETEP), granted financial resource from the Ministry of Trade, Industry & Energy, Republic of Korea (No. 20174030201790), and by Korea Electric Power Corporation. (Grant number:R18XA01).

References

[1]. R. H. Inman, H. T. C. Pedro, and C. F. M. Coimbra, “Solar forecasting methods for renewable energy integration,” Prog. Energy Combust. Sci. , vol. 39, no. 6, pp. 53 – 576, 2013.

[2]. A. Hussain, S. M. Arif, M. Aslam, Emerging renewable and sustainable energy technologies: State of the art

[3]. V. P. Singh, K. Vaibhav, and D. K. Chaturvedi, “Solar power forecasting modeling using soft computing approach,” in Engineering (NUiCONE), 2012 Nirma University International Conference on, 2012, pp. 1–5.

[4]. “The importance of wind forecasting - Renewable Energy Focus.” [Online]. Available: http://www.renewableenergyfocus.com/view/1379/the-importance-ofwind-forecasting/. [Accessed: 05-Sep-2014].

[5]. P. Poudel and B. Jang, “Solar Power Prediction Using Deep Learning Technique,” Advanced Science and Technology Letters, vol. 146, FGCN 2017, pp. 148–151, 2017.

[6]. H.S. Kim, A. Muhammad, M.S. Choi, S.J. Lee, "A study on long term solar radiation forecasting for PV in microgrid", *APAP conference Jeju Korea 2017*. P. Bacher, H. Madsen, and H. A. Nielsen, “Online short-term solar power forecasting,” Sol. Energy, vol. 83, no. 10, pp. 1772–1783, 2009

[7]. A. Heimo, “COST Action ES1002 Weather Intelligence for Renewable Energies (WIRE) CURRENT STATE Report,”August, pp. 1–164, 2012

[8]. C. Q. Y Tao H Chen, “Wind Power Prediction and Pattern Feature Based on Deep Learning Method,” pp. 1–4, 2014.

[9]. M. Khodayar and M. Teshnehlab, “Robust Deep Neural Network for Wind Speed Prediction,” 2015.

[10]. M. Aslam, J.M. Lee, S.W. Hong, S.J. Lee, E. H. Lee, “Deep Learning Application in Power System with a Case Study on Solar Irradiation Forecasting”, Conference: 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)

[11]. Bengio, P. Simard, and P. Frasconi. Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 5(2):157–166, 1994.

[12]. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio. On the properties of neural machine translation: Encoder-decoder approaches. arXiv preprint arXiv:1409.1259, 2014.