

# A deep attention-driven model to forecast solar irradiance

Abdelkader Dairi

University of Science and Technology

of Oran-Mohamed Boudiaf (USTO-MB)

Computer Science department Signal, image

and speech laboratory (SIMPA) laboratory

El Mnaouar, BP 1505, Bir El Djir 31000, Oran, Algeria

Fouzi Harrou and Ying Sun

King Abdullah University of Science and Technology (KAUST)

Computer, Electrical and Mathematical Sciences

and Engineering (CEMSE) Division

Thuwal, 23955-6900, Saudi Arabia

**Abstract**—Accurately forecasting solar irradiance is indispensable in optimally managing and designing photovoltaic systems. It enables the efficient integration of photovoltaic systems in the smart grid. This paper introduces an innovative deep attention-driven model for solar irradiance forecasting. Notably, an extended version of the variational autoencoder (VAE) is introduced by amalgamating the desirable characteristics of the bidirectional LSTM (BiLSTM) and attention mechanism with the VAE model. Specifically, the introduced approach enables the conventional VAE’s ability to model temporal dependencies by incorporating BiLSTM at the VAE’s encoder side to better extract and learn temporal dependencies embed on the solar irradiance concentration measurements. In addition, the self-attention mechanism is embedded in the VAE’s encoder side following the BiLSTM to highlight pertinent features. The performance of the proposed model is evaluated through comparisons with the recurrent neural network (RNN), gated recurrent unit (GRU), LSTM, and BiLSTM. Measurements of solar irradiance in the US and Turkey are used to evaluate the investigated models. Results confirm the superior performance of the proposed model for solar irradiance forecasting over the other models (i.e., RNN, GRU, LSTM, and BiLSTM).

**Index Terms**—Variational auto encoder, self-attention, solar irradiation ,forecasting, bidirectional recurrent neural network

## I. INTRODUCTION

The major challenge in solar energy generation consists of the volatility intermittent of PV system power production caused by weather conditions [1], [2]. The change of solar irradiation has a direct effect on the produced power by PV systems [3]. Solar irradiance is highly positively correlated with solar power harvesting, making its forecasting a relevant indicator of power generation [4], [2]. The need for accurate forecasting of solar irradiance increases with the continuing trend of growing solar power installation globally. Indeed, solar irradiance forecasting gives pertinent information to improve the power quality of electric power provided to the consumers [5]. In addition, it helps in managing energy more effectively in smart grids [6].

Effectively forecasting solar irradiance is becoming imperative to facilitate planning and managing electricity production and distribution in a smart grid. Over the last decades, various methods have been developed to enhance solar irradiance fore-

casting. Statistical models, such as Autoregressive Integrated Moving Average (ARIMA) and exponential smoothing [7], [8], are efficient in modeling a short-term linear dependence; however, more flexible models are needed for capturing nonlinear dynamics in solar irradiance. In [9], an analytical method is introduced to estimate solar irradiance, taking into account the impact of solar degradation. In [10], a hybrid Beta-kernel density estimation model is introduced for solar irradiance probability density estimation. This hybrid model showed superior prediction performance compared to the Beta distribution model. In [11], a method to forecast one-day-ahead solar irradiance based on the Markov switching model is presented for remote microgrids. Essentially, this approach utilizes available data for one-day-ahead solar irradiance forecasting to schedule energy resources in remote microgrids. In [12], a nonparametric estimator for estimating solar irradiance probability density has been introduced based on local linear regression and a root transformation approach. The authors in [13] proposed a quantile regression robustification technique to forecast solar irradiance. Based on the joint probability distribution function (PDF) of solar irradiance, in [14], a forecasting system is proposed to predict the magnitude of solar irradiance. An estimation method for solar irradiance magnitude based on PV electrical characteristics using a field-support vector regression soft sensor is proposed in [15]. However, accurate modeling and forecasting solar irradiance is a considerable challenge, requiring flexible alternatives to capture for nonlinear dynamics and time dependency.

A precise and convenient model to forecast solar irradiance can be useful for improving smart-grid management and PV systems production. Recently, data-driven deep learning models exhibited high capability in modeling time-series data in a wide range of applications, because of their flexibility and extended ability to handle complex data and capturing nonlinear dynamics [16], [17], [18], [19]. They showed successful performance in different applications, including photovoltaic power forecasting [20], anomaly detection [21], [22], [23], [24], and reinforcement learning[25], [26]. Moreover, machine learning models have been widely employed to improve solar irradiance forecasting [27], [28], [29]. In [30], a compara-

tive study between LSTM, support vector regression, principal component regression, and feed-forward neural networks (FFNN) models has been conducted to a short-term forecast of solar irradiance. Empirical results indicate the superior forecasting performance of the FFNN model compared to the other models. Moreover, It has been shown that forecast from the ensemble model based on a convex combination and quantile regression averaging (QRA) showed improved performance compared to the other models. In [31], a hybrid wavelet transformation combined with Elman Neural Network (ENN) was designed to hourly forecast Solar Irradiance in a smart grid. In [32], Long short-term memory (LSTM) has been applied for short-term solar irradiance forecasting under difficult weather circumstances. In [33], an approach using multivariate gated recurrent units (GRUs) is applied for hour-ahead solar irradiance forecasting. In [34], authors propose a Convolutional LSTM layers model to forecast solar irradiance. In prado2021flexible, a flexible data-driven approach using Convolutional Long Short-Term Memory layers is introduced for spatiotemporal forecasting of solar irradiance. The authors in [35] used a multi-model deep learning approach to enhance the forecasting accuracy of solar irradiance. This approach exploits spatial information extracted from infrared images to enhance the effectiveness of the approach. It has been shown that this approach provides improved forecasts compared to time series-based forecasts under cloudy days. In [36], a coupled model named LSTM–CNN has been employed to forecast hourly solar irradiance. The main advantage of this model consists of its ability to extract spatio-temporal features from the solar irradiance data. Results based on data from 23 locations in California demonstrated its good performance under different weather conditions.

This paper introduces an effective hybrid data-driven model to enhance the forecasting accuracy of solar irradiance. The proposed hybrid model amalgamates the desirable characteristics of BiLSTM, VAE, and the self-attention mechanism. Essentially, this paper’s main contribution consists of the extension of the traditional variational autoencoder (VAE). Towards, this end two main elements are introduced.

- The traditional VAE is based on a fully connected layer for both encoder and decoder parts. We improve the encoder efficiency by incorporating a bidirectional LSTM (BiLSTM) at the encoder side to better extract and learn temporal dependencies embed in the solar irradiance concentration measurements. This is done through forward and backward data processing, which takes advantage of future context.
- Secondly, we integrate the self-attention mechanism on the encoder side just after the BiLSTM for highlighting the pertinent features. These steps aim to enhance the input used for the variational inference process.

The proposed flexible model, the BiVAE-A, enables exploiting the advantages of BiLSTM in time-series modeling and the focus on the important features through the attention unit. We assessed the performance of the proposed model using

measurements of solar irradiance collected from two sites in the USA and Turkey. Results demonstrate the effectiveness of the model and its superior performance compared to four commonly used deep learning models (i.e., RNN, GRU, LSTM, and BiLSTM).

The remaining of this article is organized as follows. In Section II, We first present the basic idea of the proposed approach. The results and discussions were given in section III to show model performances and comparisons. The conclusions were drawn in section IV.

## II. METHODS

This study is carried out using the Bidirectional Variational Autoencoder with self-attention module, that we called BiVAE-A. The general framework of the proposed approach is summarized in Figure 1

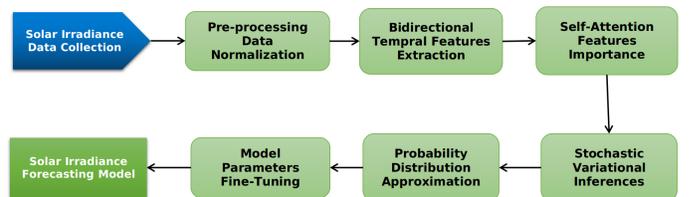


Fig. 1: Proposed approach flowchart.

### A. The proposed approach

In this paper, an effective solar irradiance magnitude forecasting framework is proposed through unsupervised learning. An extension of a generative model is introduced; namely, the Variational Autoencoder [37]. Indeed, the proposed hybrid generative model is based on bidirectional time-series data processing that aims to improve the solar irradiance magnitude forecasting quality via robust time-dependencies modeling. Figure 2 illustrates the schematic of the proposed model.

We start first with a solar irradiance data pre-processing via a normalization technique in order to uniform data range and improve the model stability. This is done by rescaling the data range to [0,1].

Indeed, a BiLSTM is used first to model time-dependencies as a temporal feature via dual processing of the input (solar irradiance). To take advantage of future context information, two LSTM models are trained simultaneously, via data processing in two directions: *forward* and *backward*, LSTM has demonstrated capabilities in different complex learning processes, such as natural language processing, speech recognition, text classification, and many others field. Indeed, LSTM is equipped with a gating mechanism, acting as a memory that enforces capturing historical data dependencies embedded in the solar irradiance time-series. The output of this step is a features space that aims to be enhanced via self-attention, which helps to highlight the relevant features [38]. Attention was designed to deal with long sequences modeling to focus attention on a special area containing features of a data

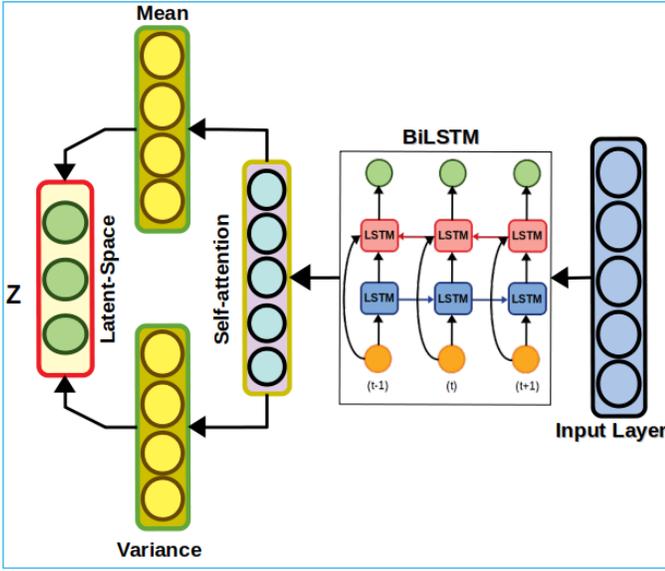


Fig. 2: Schematic representation of the proposed Bidirectional Variational Autoencoder with self-attention module.

sequence (time-series). Although a self-attention mechanism is an extension of the original attention that allows the interaction between inputs elements [39], [40], the self-attention results on a context vector (see Equation 3) based on a weighted sum of extracted features.

Let  $x_i$  denotes a specific solar irradiance and  $P(x_i)$  its prediction (See equation 1), which is a linear model.

$$P(x_i) = W_i + b \quad (1)$$

The self-attention mechanism mathematical expression is depicted as follows:

$$r_{ij} = P(x_i) \dot{P}^T(x_j) \quad (2)$$

where  $r_{ij}$  expresses the relation between  $i$ -th value and  $j$ -th value of a given sequence (input), which is used to compute the context vector  $\omega$  :

$$\omega = \text{softmax}(r_{ij}) = \frac{e^{r_{ij}}}{\sum_j e^{r_{ij}}} \quad (3)$$

The self-attention uses a regulation approach; it is anticipated to improve the forecasting quality of the solar irradiance by including the robust variation inference method with an effective regularization. The improved feature space, given the attention mechanism nonlinear transformation as a continuous representation, is more suitable for performing stochastic variational inferences [41], [37], which aim to perform an approximation of probability densities of solar irradiance via an optimization approach.

The continuous representation (the context vector  $\omega$ ) computed by the self-attention mechanism will be used to feed the covariance matrix  $\sigma$  and the mean  $\mu$  of the regularized data distributions.

Let denote  $p(x)$  the probability distribution of solar irradiance measured, VAE attempt to compute numerically  $p(x)$ , which is done through learning to draw samples similar to a data point that belongs to  $p(x)$  using a vector of latent variables represented by high-dimensional latent space  $z$ . VAE can be defined as a probabilistic model, representing a joint density  $p(z, x)$ , with  $z$  as the hidden variables, the observed variables  $x$  solar irradiance. The latent space is computed during the training as follow:

$$z = \mu + \sigma \odot \delta \quad (4)$$

With  $\delta$  a random variable :  $\delta \sim \mathcal{N}(1, 0)$ , which gives [42]:

$$z \sim \mathcal{N}(\mu, \sigma) \quad (5)$$

In fact, VAE consist of two part (deep neural network) the encoder function  $q_{\Theta}$  and the decoder  $p_{\theta}$ , with  $(\Theta, \theta)$  their parameters. The VAE objective function  $\mathcal{O}_{VAE}$  is denoted by:

$$\mathcal{O}_{VAE} = E_{q(z)}[\log p(x|z)] - KL(q(z) || p(z)) \quad (6)$$

The term  $\log p(x|z)$  expresses the log-likelihood of the observed data  $x$  (solar irradiance) for a given latent variable  $z$ . The log-likelihood similarity between samples from  $q(z)$  and the observed  $x$ . In other words, it measures the reconstruction error of the reconstruction of  $x$  from the latent space  $z$ , which is effectively the decoder mission (function). Furthermore, the *Kullback-Leibler* of  $q(z)$  and  $p(z)$  is a regularization term that measure divergence between  $(p, q)$ , which is the encoder part that sample  $z$  :  $z \sim q(z|x)$ .

A fine-tuning approach is used to adjust and optimize the pre-trained proposed model parameters (weights) to reach the best performance using a back-propagated algorithm with a supervised learning [43], and helps to attain the global minima. This step aims to enforce the learning of the mapping of a given data sequence (of solar irradiance measurement) to its next value.

Finally, the latent space is used to draw samples from the learned probability distribution of solar irradiance (the training dataset); these samples share the same features as data points used during the training; this process is called *generative*.

### III. RESULTS AND DISCUSSION

#### A. Data description

For our study, two real datasets are used to assess the forecasting accuracy of the investigated methods. The first dataset is collected in the year 2017 with 1-minute resolution in the USA, located in a parking lot canopy array monitored by the National Institute of Standard and Technology (NIST) [44]. The second one is 15 minute resolution solar irradiation data collected in the year 2018 in Izmir, Turkey.

#### B. Measurements of effectiveness

In our experiments, we adopt four metrics to evaluate forecasting accuracy: the root-mean-square error (RMSE) the

mean absolute error (MAE), the coefficient of determination ( $R^2$ ), and the explained variance (EV).

$$R^2 = \frac{\sum_{i=1}^n [(y_i - \bar{y}) \cdot (\hat{y}_i - \bar{y})]^2}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \cdot \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}}, \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}, \quad (8)$$

$$MAE = \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{n}, \quad (9)$$

$$EV = 1 - \frac{\text{Var}(\hat{\mathbf{y}} - \mathbf{y})}{\text{Var}(\mathbf{y})}, \quad (10)$$

$$MDAE(y, \hat{y}) = \text{median}(y - \hat{y}). \quad (11)$$

where here  $y$  are the actual solar irradiance values,  $\hat{y}$  are the corresponding forecasted values.

### C. Forecasting solar irradiance

In this section, we verify the performance of the proposed BiVAE-A forecasting model using solar irradiance measurements collected from the USE and Turkey. In addition, we compare the performance of the BiVAE-A approach against four well-known recurrent neural networks, namely RNN, GRU, LSTM, and BiLSTM. We assess the forecasting quality using five performance metrics: RMSE, MAE, R2, EV, and MAPE. After constructing the deep learning models using training data, we use them to forecast the future trend of solar irradiation. Forecasting results of the five considered models based on testing solar irradiance from *Turkey* and *Canopy* are shown in Figures 3 and 4, respectively. The results of the five models in Figures 3 and 4 show that forecasted irradiance is closest to the true measurement, which demonstrates their effectiveness and accuracy.

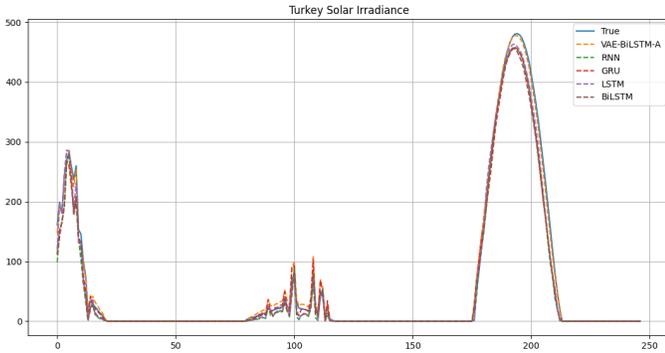


Fig. 3: solar irradiance Forecasting results of all models using Turkey Solar Irradiance

To quantitatively evaluate the performance of the proposed approach, forecasting results are evaluated using evaluation measures, namely RMSE, MAE, R2, and EV. Table I summarized the forecasting results based on Turkey data. It can be seen that the BiVAE-A scored the best R2 and EV, meaning that the presented method has caught more than 99.68% of the solar irradiance trending variation. Moreover, BiVAE-A has

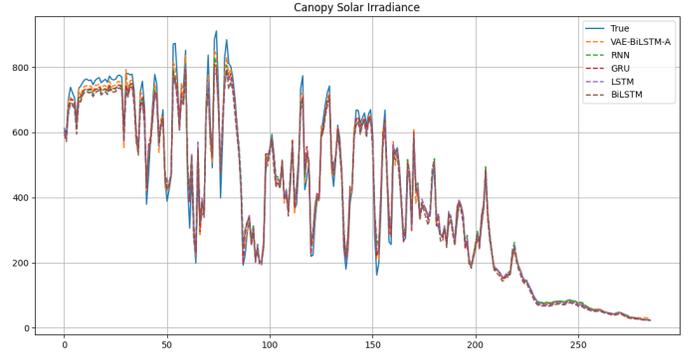


Fig. 4: solar irradiance Forecasting results of all models using Canopy Solar Irradiance

recorded the lowest modeling error  $RMSE=11.50$ ,  $MAE=4.93$ , while the other considered models recorded RMSE values between [13.92, 17.62] and MAE between with interval [7, 11.42], which demonstrates the effectiveness of the proposed approach. Thus, the BiVAE-A dominates the other model (RNN, GRU, LSTM, and BiLSTM) by achieving the best forecasting performance (Table I). This could be attributed to the high capability of the BiVAE-A in modeling nonlinear dynamics and capturing time-dependent in solar irradiance time series data.

TABLE I: solar irradiance forecasting results based on Turkey irradiance test data.

| MODEL          | RMSE            | MAE            | R2             | EV             |
|----------------|-----------------|----------------|----------------|----------------|
| RNN            | 15.287          | 7.362          | 0.994          | 0.994          |
| GRU            | 13.928          | 7.085          | 0.995          | 0.996          |
| LSTM           | 16.727          | 8.372          | 0.993          | 0.993          |
| BiLSTM         | 17.62           | 11.421         | 0.993          | 0.994          |
| <b>BiVAE-A</b> | <b>11.50383</b> | <b>4.93814</b> | <b>0.99683</b> | <b>0.99684</b> |

Similarly, statistical indicators based on *Canopy* testing data by models are listed in Table II. Results testify the superior forecasting performance of the proposed BiVAE-A model. We can observe that the proposed approach also outperforms in this series the other considered models by scoring the lowest mean error : (  $RMSE=8.3$ ,  $MAE=4.0$ ) and the highest fitting means ( $R2=99.55\%$ ) compared to the other models that recorded RMSE between [12.8, 16.7] and MAE in [5.9,7.5].

TABLE II: solar irradiance forecasting results based on Canopy solar irradiance test data.

| Model          | RMSE           | MAE            | R2             | EV             |
|----------------|----------------|----------------|----------------|----------------|
| RNN            | 16.71          | 7.572          | 0.982          | 0.985          |
| GRU            | 14.449         | 6.54           | 0.986          | 0.988          |
| LSTM           | 12.807         | 5.95           | 0.989          | 0.99           |
| BiLSTM         | 15.212         | 6.658          | 0.985          | 0.987          |
| <b>BiVAE-A</b> | <b>8.30208</b> | <b>4.07737</b> | <b>0.99551</b> | <b>0.99569</b> |

Figure 5 displays MAPE metric computed using the ob-

tained results based on *Turkey* and *Canopy* datasets. Effectively we can see that the proposed approach is recording the lowest percentage error of forecasting on both datasets. Note here that GRU and LSTM a performing better than the rest of the considered models on both datasets.

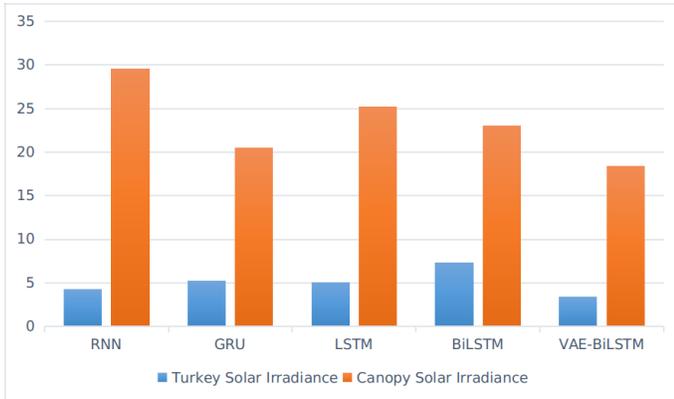


Fig. 5: solar irradiance Forecasting MAPE (%) measured for all models based on Canopy and Turkey solar irradiance

A scatter plot are used to show the forecasting accuracy of the proposed model based on Turkey and Canopy data (Figures 6). A good forecasting result means that the points cloud are forming a dense diagonal. In other words, the forecasted values coincide (fit) with real measured solar irradiance magnitude. From Figures 6, we can observe that the forecasted observation from the proposed BiVAE-A approach is forming quite a visible diagonal on Canopy and Turkey data, showing the good performance of the hybrid model again.

Overall, The obtained results demonstrate clearly that the proposed approach outperforms all the considered models; this may owe to bidirectional data processing integrated with the robust variational inference and the effectiveness of the self-attention mechanism in the learning of solar irradiance trending.

#### IV. CONCLUSION

In this paper, we introduce an effective hybrid model, called BiVAE-A, for solar irradiance forecasting. The proposed approach exploits the high ability of time-dependencies modeling of the BiLSTM combined with the self-attention mechanism to improve the quality of extracted features in the traditional VAE. The BiLSTM followed by the self-attention mechanism are integrated into the encoder part of the VAE model. The performance of the proposed hybrid model is assessed using two datasets collected from the USA and Turkey. Five metrics of effectiveness are adopted to check the forecasting accuracy of the investigated deep learning models:  $R^2$ , RMSE, MAE, EV, and MAPE. Results demonstrated that the proposed BiVAE-A model provides the best forecasting quality of solar irradiance than other deep learning models, namely RNN, GRU, LSTM, and BiLSTM. Of course, this study demonstrated the promising performances of deep learning models in forecasting solar irradiance time series without exogenous information.

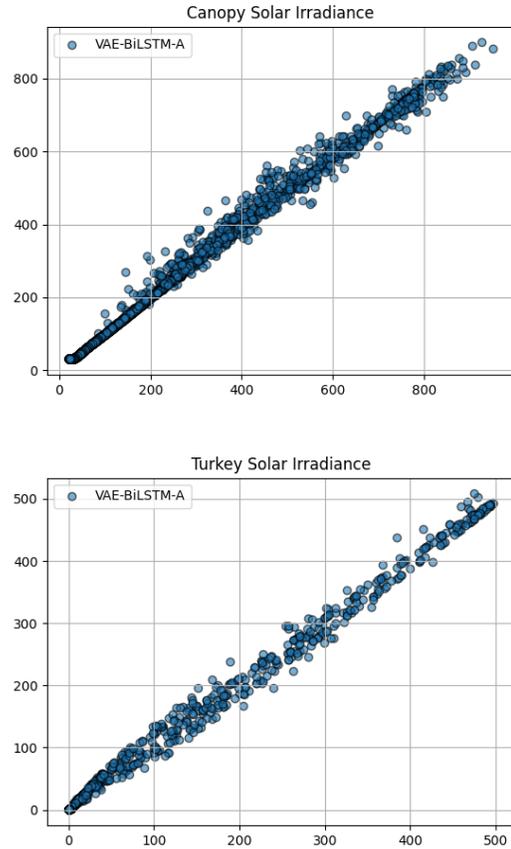


Fig. 6: Solar irradiance forecasting of BiVAE-A based on Turkey and Canopy datasets.

#### REFERENCES

- [1] A. Dairi, F. Harrou, Y. Sun, and S. Khadraoui, "Short-term forecasting of photovoltaic solar power production using variational auto-encoder driven deep learning approach," *Applied Sciences*, vol. 10, no. 23, p. 8400, 2020.
- [2] J. Lee, W. Wang, F. Harrou, and Y. Sun, "Reliable solar irradiance prediction using ensemble learning-based models: A comparative study," *Energy Conversion and Management*, vol. 208, p. 112582, 2020.
- [3] Y. Feng, W. Hao, H. Li, N. Cui, D. Gong, and L. Gao, "Machine learning models to quantify and map daily global solar radiation and photovoltaic power," *Renewable and Sustainable Energy Reviews*, vol. 118, p. 109393, 2020.
- [4] Y. El Mghouchi, E. Chham, M. Krikiz, T. Ajzoul, and A. El Bouardi, "On the prediction of the daily global solar radiation intensity on south-facing plane surfaces inclined at varying angles," *Energy conversion and management*, vol. 120, pp. 397–411, 2016.
- [5] H.-T. Yang, C.-M. Huang, Y.-C. Huang, and Y.-S. Pai, "A weather-based hybrid method for 1-day ahead hourly forecasting of pv power output," *IEEE transactions on sustainable energy*, vol. 5, no. 3, pp. 917–926, 2014.
- [6] D. Yang, C. Gu, Z. Dong, P. Jirutitjaroen, N. Chen, and W. M. Walsh, "Solar irradiance forecasting using spatial-temporal covariance structures and time-forward kriging," *Renewable Energy*, vol. 60, pp. 235–245, 2013.
- [7] Z. Dong, D. Yang, T. Reindl, and W. M. Walsh, "Short-term solar irradiance forecasting using exponential smoothing state space model," *Energy*, vol. 55, pp. 1104–1113, 2013.
- [8] M. H. Alsharif, M. K. Younes, and J. Kim, "Time series ARIMA model for prediction of daily and monthly average global solar radiation: The

- case study of Seoul, South Korea,” *Symmetry*, vol. 11, no. 2, p. 240, 2019.
- [9] S. Y. Alsadi and Y. F. Nassar, “Estimation of solar irradiance on solar fields: An analytical approach and experimental results,” *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1601–1608, 2017.
  - [10] M. Wabwah, T. H. M. EL-Fouly, B. Zahawi, and S. Feng, “Hybrid beta-ke model for solar irradiance probability density estimation,” *IEEE Transactions on Sustainable Energy*, vol. 11, no. 2, pp. 1110–1113, 2020.
  - [11] A. Shakya, S. Michael, C. Saunders, D. Armstrong, P. Pandey, S. Chalise, and R. Tonkoski, “Solar irradiance forecasting in remote microgrids using markov switching model,” *IEEE Transactions on Sustainable Energy*, vol. 8, no. 3, pp. 895–905, 2017.
  - [12] M. Wabwah, S. Feng, T. H. M. EL-Fouly, and B. Zahawi, “Root-transformed local linear regression for solar irradiance probability density estimation,” *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 652–661, 2020.
  - [13] E. Ranganai and C. Sigauke, “Capturing long-range dependence and harmonic phenomena in 24-hour solar irradiance forecasting: A quantile regression robustification via forecasts combination approach,” *IEEE Access*, vol. 8, pp. 172 204–172 218, 2020.
  - [14] M. Kakimoto, Y. Endoh, H. Shin, R. Ikeda, and H. Kusaka, “Probabilistic solar irradiance forecasting by conditioning joint probability method and its application to electric power trading,” *IEEE Transactions on Sustainable Energy*, vol. 10, no. 2, pp. 983–993, 2019.
  - [15] J. Ma, H. Jiang, K. Huang, Z. Bi, and K. L. Man, “Novel field-support vector regression-based soft sensor for accurate estimation of solar irradiance,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 12, pp. 3183–3191, 2017.
  - [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.
  - [17] A. Graves, A. . rahman Mohamed, and G. E. Hinton, “Speech recognition with deep recurrent neural networks,” *CoRR*, vol. abs / 1303.5778, 2013. [Online]. Available: <http://arxiv.org/abs/1303.5778>
  - [18] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, p. 1735–1780, Nov. 1997. [Online]. Available: <https://doi.org/10.1162/neco.1997.9.8.1735>
  - [19] F. Harrou, Y. Sun, A. S. Hering, M. Madakyaru *et al.*, *Statistical Process Monitoring Using Advanced Data-Driven and Deep Learning Approaches: Theory and Practical Applications*. Elsevier, 2020.
  - [20] F. Harrou, F. Kadri, and Y. Sun, “Forecasting of photovoltaic solar power production using lstm approach,” in *Advanced Statistical Modeling, Forecasting, and Fault Detection in Renewable Energy Systems*. IntechOpen, 2020.
  - [21] F. Harrou, M. M. Hittawe, Y. Sun, and O. Beya, “Malicious attacks detection in crowded areas using deep learning-based approach,” *IEEE Instrumentation & Measurement Magazine*, vol. 23, no. 5, pp. 57–62, 2020.
  - [22] W. Wang, J. Lee, F. Harrou, and Y. Sun, “Early detection of parkinson’s disease using deep learning and machine learning,” *IEEE Access*, vol. 8, pp. 147 635–147 646, 2020.
  - [23] A. Dairi, T. Cheng, F. Harrou, Y. Sun, and T. Leiknes, “Deep learning approach for sustainable wwtpp operation: A case study on data-driven influent conditions monitoring,” *Sustainable Cities and Society*, vol. 50, p. 101670, 2019.
  - [24] F. Harrou, A. Dairi, Y. Sun, and F. Kadri, “Detecting abnormal ozone measurements with a deep learning-based strategy,” *IEEE Sensors Journal*, vol. 18, no. 17, pp. 7222–7232, 2018.
  - [25] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, “Mastering the game of go with deep neural networks and tree search,” 2016.
  - [26] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis.
  - [27] F. Wang, Y. Yu, Z. Zhang, J. Li, Z. Zhen, and K. Li, “Wavelet decomposition and convolutional lstm networks based improved deep learning model for solar irradiance forecasting,” *applied sciences*, vol. 8, no. 8, p. 1286, 2018.
  - [28] H. Zang, L. Liu, L. Sun, L. Cheng, Z. Wei, and G. Sun, “Short-term global horizontal irradiance forecasting based on a hybrid cnn-lstm model with spatiotemporal correlations,” *Renewable Energy*, vol. 160, pp. 26–41, 2020.
  - [29] G. Moreno, P. Martin, C. Santos, F. J. Rodríguez, and E. Santiso, “A day-ahead irradiance forecasting strategy for the integration of photovoltaic systems in virtual power plants,” *IEEE Access*, vol. 8, pp. 204 226–204 240, 2020.
  - [30] T. Mutavhatsindi, C. Sigauke, and R. Mbuva, “Forecasting hourly global horizontal solar irradiance in south africa using machine learning models,” *IEEE Access*, vol. 8, pp. 198 872–198 885, 2020.
  - [31] X. Huang, J. Shi, B. Gao, Y. Tai, Z. Chen, and J. Zhang, “Forecasting hourly solar irradiance using hybrid wavelet transformation and elman model in smart grid,” *IEEE Access*, vol. 7, pp. 139 909–139 923, 2019.
  - [32] Y. Yu, J. Cao, and J. Zhu, “An lstm short-term solar irradiance forecasting under complicated weather conditions,” *IEEE Access*, vol. 7, pp. 145 651–145 666, 2019.
  - [33] J. Wojtkiewicz, M. Hosseini, R. Gottumukkala, and T. L. Chambers, “Hour-ahead solar irradiance forecasting using multivariate gated recurrent units,” *Energies*, vol. 12, no. 21, p. 4055, 2019.
  - [34] I. I. Prado-Rujas, A. García-Dopico, E. Serrano, and M. S. Pérez, “A flexible and robust deep learning-based system for solar irradiance forecasting,” *IEEE Access*, vol. 9, pp. 12 348–12 361, 2021.
  - [35] M. Ajith and M. Martínez-Ramón, “Deep learning based solar radiation micro forecast by fusion of infrared cloud images and radiation data,” *Applied Energy*, vol. 294, p. 117014, 2021.
  - [36] P. Kumari and D. Toshniwal, “Long short term memory–convolutional neural network based deep hybrid approach for solar irradiance forecasting,” *Applied Energy*, vol. 295, p. 117061, 2021.
  - [37] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” *stat*, vol. 1050, p. 1, 2014.
  - [38] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio, “Show, attend and tell: Neural image caption generation with visual attention,” in *International conference on machine learning*, 2015, pp. 2048–2057.
  - [39] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” in *Advances in neural information processing systems*, 2017, pp. 5998–6008.
  - [40] A. Dairi, F. Harrou, S. Khadraoui, and Y. Sun, “Integrated multiple directed attention-based deep learning for improved air pollution forecasting,” *IEEE Transactions on Instrumentation and Measurement*, 2021.
  - [41] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, “Variational inference: A review for statisticians,” *Journal of the American statistical Association*, vol. 112, no. 518, pp. 859–877, 2017.
  - [42] I. Drori, *Deep Variational Inference*. Cham: Springer International Publishing, 2020, pp. 361–376.
  - [43] G. E. Hinton, “A practical guide to training restricted boltzmann machines,” in *Neural networks: Tricks of the trade*. Springer, 2012, pp. 599–619.
  - [44] M. Boyd, “Performance data from the nist photovoltaic arrays and weather station,” *J. Res. Natl. Inst. Stand. Technol.*, vol. 122, no. 40, p. 40, 2017.