Application of wavelet and seasonal-based emotional ANN (EANN) models to predict solar irradiance

Energy Reports 12 (2024) 3258–3277

Contents lists available at ScienceDirect

Energy Reports

journal homepage: www.elsevier.com/locate/egyr

Application of wavelet and seasonal-based emotional ANN (EANN) models to predict solar irradiance

Vahid Nourani [a,b], Nazanin Behfar [c,*], Anne Ng [d], Chunwei Zhang [e], Fahreddin Sadikoglu [f]

[a] Center of Excellence in Hydroinformatics and Faculty of Civil Engineering, University of Tabriz, 29 Bahman Ave., Tabriz, Iran
[b] Faculty of Civil and Environmental Engineering, Near East University, Nicosia, via Mersin 10, Turkey
[c] Department of Water Engineering and Management, Faculty of Engineering Technology, University of Twente, Enschede 7500 AE, the Netherlands
[d] College of Engineering, Information Technology and Environment, Charles Darwin University, Ellengowan, Brinkin, NT 0810, Australia
[e] Multidisciplinary Center for Infrastructure Engineering, Shenyang University of Technology, Shenyang 110870, China
[f] Artificial Intelligence Center, Odlar Yurdu University, Koroglu Rahimov Street, 13, Baku AZ1072, Azerbaijan

ARTICLE INFO

Keywords:
Solar irradiance modeling
Emotional artificial neural network (EANN)
Seasonal model
Wavelet transform
Iran
USA

ABSTRACT

This study models solar irradiance at six stations in Iran and the USA on an hourly scale. We explored two seasonal emotional artificial neural networks (EANN): sequence-EANN (SEANN) and wavelet EANN (WEANN). Analyzing ten years of climatic and solar data, we evaluated uncertainty using prediction intervals (PIs) computed via the bootstrap method based on artificial neural networks (ANNs). Unlike standalone EANNs, the proposed seasonal models effectively captured seasonal information and leveraged time series processing advantages. Utilizing Wavelet and Fourier transforms, these models captured long-short autoregressive dependencies in solar irradiance, addressing extended seasonal dependencies. Results showed that the seasonal EANN models outperformed the classic EANN model by approximately 15 % and the classic feed-forward neural network (FFNN) by about 25 % in both training and testing. The WEANN model demonstrated the highest performance in PIs, with an average normalized mean PI width (NMPIW) of 0.8 and an average PI coverage probability (PICP) of 0.96.

1. Introduction

Climate change and global warming are escalating due to the exploitation of fossil fuel resources. Consequently, the adoption of renewable energy sources has become a pressing concern. While renewable energy sources offer numerous benefits, such as environmental sustainability and an essentially limitless supply, they face a significant challenge—high intermittency (Mirparizi et al., 2023). The output power of renewable sources fluctuates, resulting in time series data with a variable non-stationary nature. The introduction of solar power comes with challenges related to the solar resource. Ground irradiance variability, for instance, makes monitoring and storing power demanding and economically inefficient. Solar generation, being more uncertain than traditional fossil energy sources, requires increased control and storage capacity to manage ancillary service conditions. Volatile and extensive changes in irradiance, often referred to as "ramp events," are typically influenced by substantial and widespread cloud cover (Pfister et al., 2003). Addressing variations in electrical supply and power quality reservation, ramp events necessitate ancillary services to ramp up and down. Accurate solar projection reduces uncertainty in solar generation, enhancing the cost-effectiveness of solar applications (Yang et al., 2014). For this purpose, modeling solar radiation in diverse geographic areas with specific solar potential plays a crucial role in obtaining precise information and reducing uncertainty in accessing clean energy (Akarslan et al., 2018). Various methods are employed for solar radiation forecasting, with Artificial Intelligence (AI) methods gaining extensive use in recent decades (e.g., Yadav and Chandel, 2014; Piri and Kisi, 2015; Sharifi et al., 2016; Gutierrez-Corea et al., 2016; Chen and Kartini, 2017; Meenal and Selvakumar, 2018; Nourani et al., 2019; Koo et al., 2020; Tao et al., 2021; Geetha et al., 2022; Sharifi et al., 2022; Nourani et al., 2022; Abdallah et al., 2023; Ehteram and Shabanian, 2023). Despite the efficacy of AIs in solar irradiance modeling, challenges may arise when there is insufficient data for training the AI model or when the data exhibit pronounced non-stationarity with seasonal fluctuations. In such cases, the modeling process may struggle to detect seasonal trends in time series, which are crucial components of

* Corresponding author.
E-mail address: n.behfar@utwente.nl (N. Behfar).

https://doi.org/10.1016/j.egyr.2024.09.011
Received 14 February 2024; Received in revised form 14 July 2024; Accepted 4 September 2024
Available online 13 September 2024

natural phenomena. The wavelet transform (WT), a powerful mathematical method, provides time-frequency characteristics of a signal during a given time period (Dabuechies, 1990). In recent years, WT has found widespread use in investigating time series fluctuations, cycles, and trends. Moreover, WT-equipped hybrid models have been introduced to enhance prediction accuracy (Nourani et al., 2014), and particularly in solar irradiance modeling (e.g., Zhang et al., 2011; Yadav and Chandel, 2014; Khaled et al., 2022). Lyu et al. (2014) introduced an innovative approach for assessing solar irradiance. They employed WT as a data preprocessing technique to denoise meteorological time series. This research constructed predictive models using Artificial Neural Networks (ANN) and support vector machines for various cities in California, Kentucky, and New York, utilizing Global Horizontal Irradiance (GHI). The results demonstrated increased efficiency, suggesting the method potential as a generalized approach. Sharma et al. (2016) proposed a hybrid Wavelet Neural Network (WNN) for short-term solar irradiance projection. Morlet and Mexican hat wavelets were combined as activation functions for hidden-layer neurons in WNNs. Comparative analysis indicated that WNN modeling outperformed other methods in predicting future variations of solar irradiance, demonstrating its effectiveness. Dewangan et al. (2017) implemented WNN with Levenberg-Marquardt training for solar irradiance forecasting to detect solar power output. The study concluded that this methodology could be more efficient in general applications and more precise than regular sigmoidal neural networks (SNN). Results showed that the model's execution is straightforward and could improve the forecasting skills.

Recently a unique type of ANN model, known as emotional ANN (EANN) model, has been proposed by integrating artificial emotions with the classic ANN framework (Khashman, 2008; Lotfi and Akbarzadeh-T., 2016). Drawing inspiration from biological processes, where hormonal activities impact the neuropsychological responses of animals, the proposed EANN model incorporates a feedback loop between hormones and neurons. This biological theory suggests that such a loop could enhance the learning capacity of the network (Nourani, 2017). While the concept of EANN has been explored and applied in various engineering domains (e.g., Babaie et al., 2008; Pavlos et al., 2011; Abdi et al., 2012; Nourani, 2017; Reddy et al., 2021), its application within the context of energy modeling, specifically in solar irradiance modeling, is novel and has not been extensively discussed. In this study, two newly developed EANN models are explored and applied for solar irradiance modeling: sequence-EANN (SEANN) and wavelet-EANN (WEANN). The EANN model is calibrated using a seasonal sequence (SS) of solar irradiance time series to accurately identify shorter and longer seasonal patterns. Consequently, in the developed WEANN model, WT is applied initially to decompose time series, and then the dominant sub-series are fed into the EANN model to predict solar irradiance.

Despite the precision achieved by AI models, they frequently lack information regarding uncertainty. Conventional AI techniques, which provide singular outputs, fail to account for sampling error or uncertainty. To tackle this limitation, the incorporation of confidence intervals (CIs) and prediction intervals (PIs) can be instrumental in estimating uncertainty in irradiance modeling, as demonstrated by studies such as Torregrossa et al. (2016). PIs, encompassing both aleatoric and epistemic uncertainties, are regarded as more informative than CIs, as highlighted by Khosravi et al. (2011). The commonly employed bootstrap method, along with its various adaptations like Moving Blocks Bootstrap (MBB), is a prevalent approach for estimating PIs in ANN-based modeling and it was employed in this study to verify the uncertainty of the developed EANN based models for prediction of solar irradiance as the first application in the field of solar irradiance modeling.

In summary, despite significant advancements in solar irradiance modeling, several gaps remain in the existing literature. Traditional ANN models often fail to capture the seasonal variations inherent in solar irradiance data, leading to less accurate predictions. Additionally, while various studies have employed ANN models for solar irradiance prediction, few have addressed the need for robust uncertainty quantification in their predictions. This gap is critical because understanding the uncertainty in predictions is essential for the reliable integration of solar energy into power grids. Recent studies highlight the challenges and advancements in solar irradiance prediction. For instance, Tercha et al. (2024) review machine learning techniques for forecasting temperature and solar irradiance, emphasizing the need for accurate predictions to optimize photovoltaic systems. Furthermore, recent research by Wang et al. (2024) underscores the critical role of uncertainty quantification in solar irradiance modeling. Their study demonstrates that accurate estimation of prediction uncertainties enhances decision-making processes in energy management, ensuring more reliable grid operations and improved economic outcomes. Our study aims to fill these gaps by introducing seasonal EANN that not only incorporate seasonal dependencies into the modeling process but also compute PIs to quantify the uncertainty in solar irradiance predictions. This dual approach of enhancing prediction accuracy and providing a measure of prediction reliability represents a significant step forward in the field of solar irradiance modeling. Therefore, the novelty of this study lies in the application of seasonal EANN to predict solar irradiance and compute PIs. Unlike traditional ANN models, which often overlook seasonal patterns and dependencies, our approach integrates seasonal variations into the modeling process. This integration enhances the model's ability to capture and predict the intricate fluctuations of solar irradiance over time. Additionally, the use of EANNs for calculating PIs provides a robust measure of uncertainty in predictions, offering a comprehensive understanding of model reliability. This dual focus on seasonal prediction and uncertainty quantification represents a significant advancement in solar irradiance modeling, setting our work apart from conventional ANN methodologies. The importance of such modeling in the collection and consumption of solar energy cannot be overstated.

The objectives of this study could be listed as: i) To implement seasonal-based EANN methods for solar irradiance simulation, utilizing data from six stations situated in Iran and the USA to assess the modeling at different regions. ii) To assess and compare the performance of the introduced seasonal-based EANN models against conventional models. iii) To enhance the practicality of the models by incorporating PIs and provide decision-makers with a dependable tool for assessing potential solar irradiance values and associated uncertainties.

2. Materials and methods

This section introduces the study area and data, then outlines the concepts of traditional EANN and FFNN. It then introduces the structure of two proposed seasonal EANN methods (refer to Fig. 1) and finally explains the MMB procedure.

2.1. Study area and data

To model solar irradiance, two countries with diverse stations situated at different latitudes were chosen to assess the impact of geographic location on modeling accuracy. Meteorological hourly data were gathered from six stations in Iran (Tabriz, Rasht, and Zahedan stations) and the USA (Tucson, Minneapolis, and Miami stations) over the ten-year period from January 2011 through May 2021 (see Fig. 2). These locations were strategically selected for their diverse latitudes, resulting in varying patterns of solar irradiance and climatic conditions, aimed at capturing a wide range of solar radiation intensities and seasonal variations. Specifically chosen to encompass a broad spectrum of solar irradiance variability influenced by geographical positioning and local weather patterns, these sites enable us to effectively explore the performance and applicability of our seasonal EANN models across a range of climatic regimes.

Iran exhibits a wide-ranging climate, varying from subtropical along the Caspian Sea coast and in the northern forests to continental and arid on the Iranian Plateau. The northwestern Iranian city of Tabriz

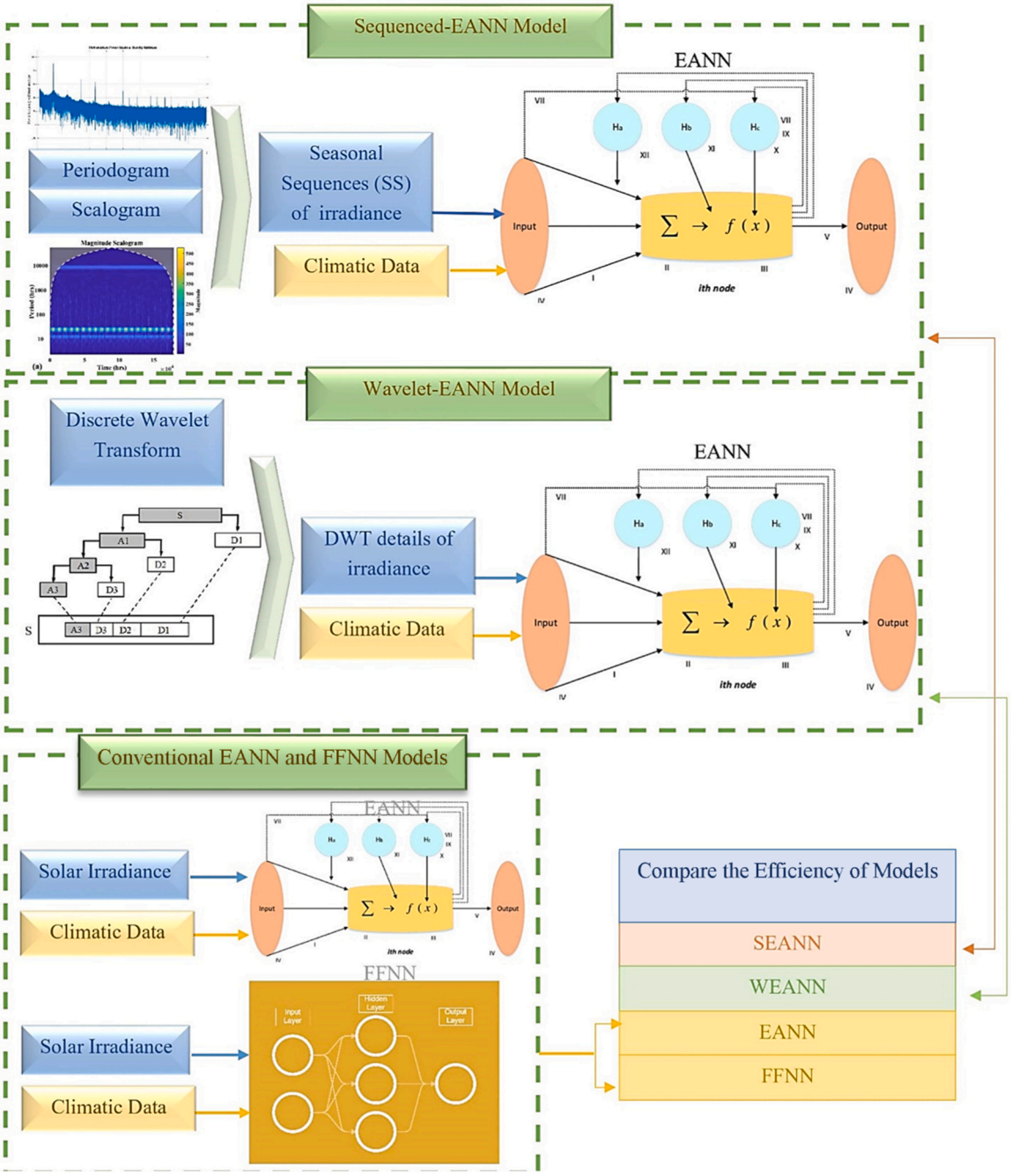


Fig. 1. Schematic of proposed methodology for the solar irradiance modeling employing seasonal-based EANN models, including SEANN and WEANN.

(38.08°N) experiences a continental climate with distinct seasons, bordering on a cold semi-arid climate. With approximately 2750 hours of sunshine annually, Tabriz undergoes a mild spring, a dry and semi-hot summer, a humid and rainy fall, and a snowy cold winter. Situated in the northern part of Iran, Rasht City (37.28°N) features a humid subtropical climate characterized by hot summers and cold (sometimes snowy) winters. Due to frequent rainfall, Rasht station records a minimum of sunny hours, approximately 1520 hours per year. Zahedan (29.49°N), located in southeastern Iran, is renowned for its extended sunshine hours, reaching around 3160 per year. The climate in this city is categorized by long, hot, and arid summers, accompanied by cold and dry winters. Zahedan seldom experiences precipitation, and the sky tends to remain mostly clear.

The United States exhibits a diverse range of climate types and geographical features, including mountains, deserts, and oceans, reflecting the variation in latitude across the country. Miami (25.76°N), situated in southeastern USA, has a subtropical climate characterized by humidity and warmth in summer, as well as mild and short winters. With an annual average of approximately 3150 sun hours, the city receives a significant amount of sunlight throughout the year. The sunniest months are March and April, while the lowest levels of sunshine occur in winter. Tucson (32.22°N), located in the southwest USA, experiences two predominant seasons—scorching summers and dry-cold winters. Fall and spring are characterized by sunny and dry climate conditions. The city boasts an estimated 3850 hours of sunshine annually, making it one of the sunniest places in the country. The average sunshine amount is 96 % throughout the year, with the maximum measured at 92–94 % from April to June. The lowest proportion of sunshine, 79 %, occurs in December during the North American Monsoon. Minneapolis (44.97°N), located in the northern USA, holds the distinction of having the lowest

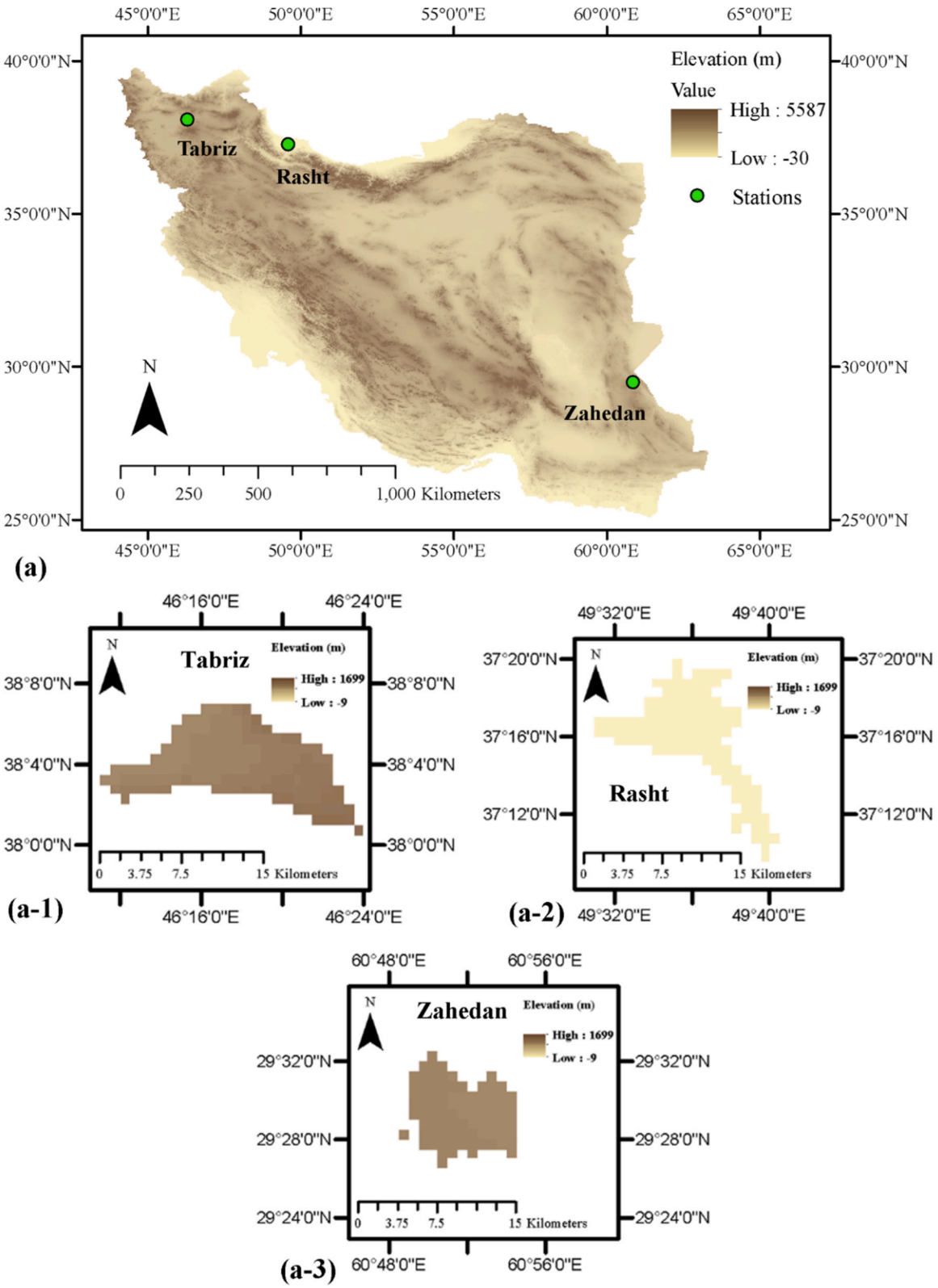


Fig. 2. Study areas and selected stations, 3 stations from **(a)** Iran, **(a-1)** Tabriz, **(a-2)** Rasht, **(a-3)** Zahedan and 3 stations from **(b)** USA, **(b-1)** Tucson, **(b-2)** Minneapolis, **(b-3)** Miami.

average temperature among large cities. The city experiences cold winters and hot, humid summers, reflecting significant seasonal variations in weather. Interestingly, despite the cold winters, Minneapolis sees the highest levels of sunshine in mid-winter compared to other parts of the country. On an annual basis, an average of 2170 hours of sunshine is calculated for the city. Due to its northern location in the US, Minneapolis receives fewer hours of sunshine compared to the other stations.

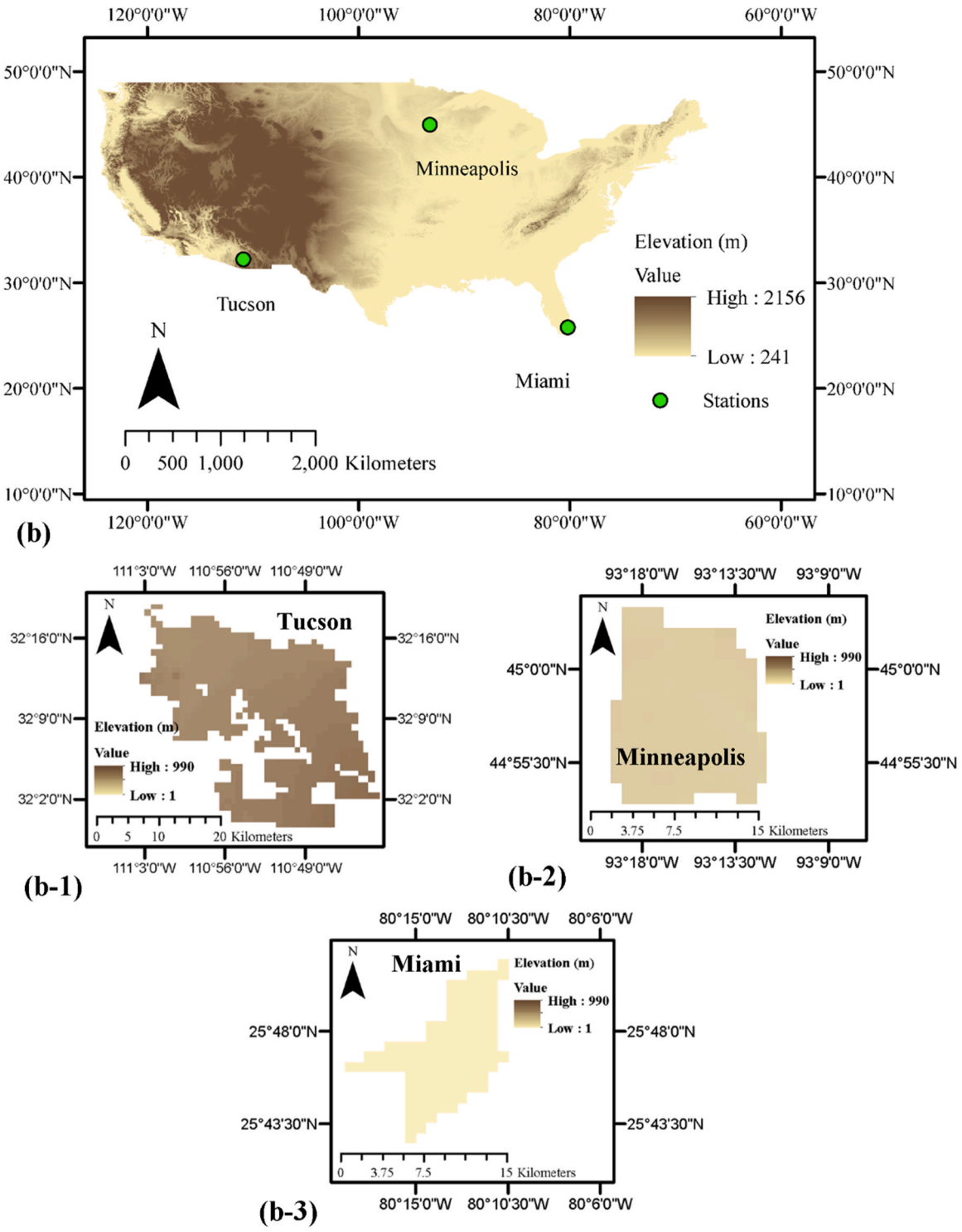


Fig. 2. (*continued*).

The data used in this study were sourced from the National Aeronautics and Space Administration (NASA) at https://power.larc.nasa.gov. The dataset includes ten years (2011–2021) information on solar irradiance (I) measured in watt-hours per square meter (Wh/m2) and temperature (T) recorded at 2 m above the surface of the Earth, expressed in degrees Celsius (C). For reference, Table 1 provides the statistical summary of the irradiance and temperature data from the stations involved in this study.

In analyzing the dataset sourced from NASA, it is important to acknowledge potential influences on data quality and representativeness. The satellite-based nature of the data means it may be subject to weather conditions affecting measurements, particularly cloudiness, which can obscure solar irradiance readings. Additionally, since the focus is on solar irradiance, values are naturally zero during nighttime periods. While there were no gaps identified in the dataset itself, variations in satellite coverage, calibration procedures, and local weather patterns, including cloud cover, could still introduce biases. Each year and station's representation may vary due to such factors, impacting the dataset's overall consistency. Nonetheless, efforts were made to select stations across Iran and the USA that provide a diverse

Table 1
Statistics of the used hourly data of the stations for solar irradiance predictions.

Station	Location	Coordinates	Parameter (Unit)	All data				Training data				Verification data			
				Min	Max	Mean	Standard deviation	Min	Max	Mean	Standard deviation	Min	Max	Mean	Standard deviation
Tabriz	Latitude	38°04′47″ N	I (Wh/m^2)	0.0	1030.4	316.3	290.0	0.0	1023.6	315.6	288.8	0.0	1030.4	316.8	292.2
	Longitude	46°17′30″ E	T (C)	−21.0	38.8	11.6	11.5	−21.0	38.8	11.5	11.5	−17.2	38.6	11.5	11.6
	Altitude	1396 m													
Rasht	Latitude	37°16′33″ N	I (Wh/m^2)	0.0	972.4	256.8	256.0	0.0	966.6	255.9	254.1	0.0	972.4	258.7	260.7
	Longitude	49°35′19″ E													
	Altitude	5 m	T (C)	−4.2	37.1	17.7	7.8	−4.2	36.6	17.7	7.8	−1.6	37.1	17.8	7.8
Zahedan	Latitude	29°29′46″ N	I (Wh/m^2)	0.0	1083.9	402.4	336.7	0.0	1083.9	403.2	337.0	0.0	1064.7	399.5	335.6
	Longitude	60°51′46″ E													
	Altitude	1380 m	T (C)	−10.7	41.9	20.5	10.0	−10.7	41.1	20.5	10.1	−7.4	41.9	20.4	10.0
Miami	Latitude	25°46′27″ N	I (Wh/m^2)	0.0	1022.9	336.3	286.9	0.0	1022.9	336.1	287.1	0.0	1011.6	337.9	286.8
	Longitude	80°11′37″ W													
	Altitude	25 m	T (C)	5.9	31.9	25.7	3.3	5.9	31.8	25.6	3.4	9.3	31.9	26.0	3.1
Tucson	Latitude	32°13′18″ N	I (Wh/m^2)	0.0	1084.7	388.0	328.8	0.0	1084.7	389.6	329.3	0.0	1075.8	384.7	327.8
	Longitude	110°55′35″ W													
	Altitude	757 m	T (C)	−8.1	44.8	22.0	9.9	−8.1	44.7	21.9	9.9	−4.6	44.8	22.3	9.7
Minneapolis	Latitude	44°58′47″ N	I (Wh/m^2)	0.0	993.1	258.6	258.7	0.0	993.1	260.8	260.2	0.0	979.2	251.9	254.4
	Longitude	93°15′49″ W													
	Altitude	262 m	T (C)	−38.7	41.6	8.2	13.7	−38.7	41.6	8.4	13.8	−37.5	34.5	7.5	13.5

representation of climatic conditions, aiming to mitigate potential biases and enhance the study's robustness. By acknowledging these considerations, we provide a clearer context for interpreting the findings of our study on the performance of seasonal EANN models in predicting solar irradiance.

It is crucial to note that the dataset was split into calibration and verification tasks, with proportions designated as 75 % for calibration and 25 % for verification. Additionally, normalization was applied to the datasets to enhance the calibration process, ensuring greater precision (Nourani, 2017):

$$S_{norm} = \frac{S_{(t)} - S_{\min(t)}}{S_{\max(t)} - S_{\min(t)}} \quad \leq 1 \tag{1}$$

where S_{norm} is the normalized value of the $S_{(t)}$; $S_{\max(t)}$ and $S_{\min(t)}$ are the maximum and minimum values of the observed data, respectively.

2.2. Structure of seasonal EANN models

The EANN represents an advanced version of traditional ANNs by integrating emotions to modify each neuron's response. EANNs employ a feedback loop between hormones and neurons, illustrated in Fig. 3, to enhance learning. Inspired by biological systems, this emotional component allows EANNs to dynamically adjust activation thresholds and synaptic weights based on emotional states. Positive emotions, for instance, may lower thresholds, facilitating faster learning and adaptation, whereas negative emotions may raise them, promoting cautious decision-making and robust error learning. This emotional feedback mechanism in EANN mirrors neurobiological processes observed in living organisms, where emotions significantly influence behavior and cognitive functions. By integrating emotional dynamics, EANN not only enhances neural network adaptability and learning efficiency but also provides a biologically plausible framework for modeling complex cognitive processes, offering a promising avenue for advancing AI research. In EANNs, hormones act as dynamic coefficients initialized based on input and target data patterns. Throughout training iterations, these hormones dynamically adapt, influencing weights, network function, and activation. According to Fig. 3, the output Y_i of the i th neuron in an EANN with three hormonal glands of H_a, H_b, and H_c is calculated as:

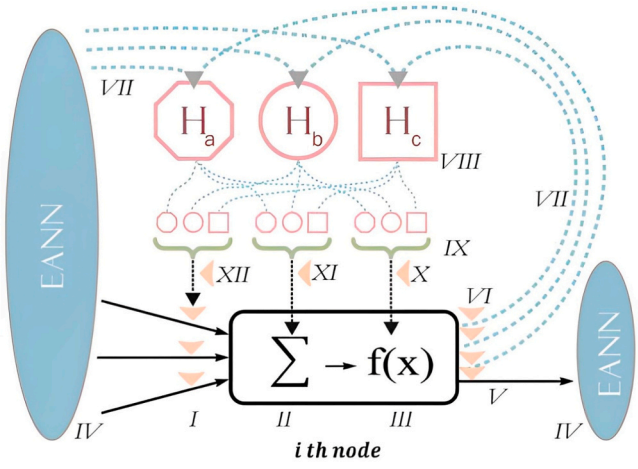


Fig. 3. A node of EANN and emotional unit (I = weights applied to inputs, II = net function, III = activation function, IV = input/output unit, VI = glandity, VIII = hormonal unit, hormones Hh, IX = hormones net unit; V, VII, X, XI, XII = respectively weights applied on net output, hormones from input/output, activation function, net function, and input static weights).

$$Y_i = \left(\lambda_i + \sum_h \sigma_{i.h} H_h\right) \times f\left(\sum_j \left[(\beta_i + \sum_h \zeta_{i.h} H_h) \times (\theta_{i.j} + \sum_h \phi_{i.j.k} H_h) X_{i.j}\right) + (\alpha_i + \sum_h x_{i.h} H_h)\right]\Bigg) \tag{2}$$

here the overall hormonal level H of the EANN is computed as:

$$H_h = \sum_i H_{i.h} \quad (h = a.b.c) \tag{3}$$

In Eq. 2, term (1) represents the weight applied to the activation function ($f()$), comprising both the constant neural weight λi and the dynamic hormonal weight $\sum_h \sigma_{i.h} H_h$. Term (2) denotes the weight applied to the net function (summation), term (3) applies weights to input values $x_{i.j}$ from the j th neuron of the previous layer, and term (4) represents the bias of the net function, which includes neural and hormonal weights a_i and $\sum_h x_{i.h} H_h$. The contribution of the system's hormonal level (H_h) in each hormonal weight is controlled by parameters $\sigma_{i.h}$, $\zeta_{i.h}$, $\phi_{i.h}$ and $X_{i.h}$. The output Y_i of the i th neuron provides hormonal feedback $H_{i.j}$ to the system as:

$$H_{i.h} = glandity_{i.h} \times Y_i \tag{4}$$

where $glandity_{i.h}$ serves as a calibration parameter producing the hormonal level of each gland. Various schemes can initialize the hormone values H_h, considering input patterns such as the mean of input parameter values for each sample (input vector). Subsequently, through the training process, the hormone values are updated to achieve reliable agreement between observed and computed target values. Simplified versions of the EANN can be proposed by omitting certain modules and components of this general network (Thenius et al., 2013). For further details on EANN structure and equations, refer to Nourani (2017) and Thenius et al. (2013).

According to the schematic presented in Fig. 1, the solar irradiance modeling in this study was conducted using two approaches. In the first approach, named SEANN modeling, effective periodicities were extracted from the irradiance time series using the periodogram (by Fourier transform) and scalogram (by Continuous Wavelet Transform (CWT)). Subsequently, the extracted SS irradiance time series, along with climatic variables, were inputted into the EANN to predict solar irradiance. The proposed SEANN was designed to detect the seasonality dependency within the time series. In the SEANN model temperature and cloud cover as climate variables and SS time series were considered as input sets, where SS is formulated as:

$$SS_t = (I_{(t-n)-k\times i}, \quad \cdots, I_{(t-1)-k\times i}, I_{t-k\times i}) \tag{5}$$

in which I shows the input dataset, t and n correspond to time step and lag window, respectively. The time step between SS and output is represented by i. Given that the primary frequency in the signals is on an hourly time scale, "i" takes on the value of 1 for one day ago and one year ago. Additionally, k corresponds to 24 hours for one day, and 8760 hours for a year. The SS dataset reveals the seasonal features of solar irradiance, encapsulating the long-term properties of the process. While the EANN struggles to capture temporal information (seasonality), SS enhances the model's ability to identify very long features, resulting in more effective assessment of solar irradiance. SS serves as the model input, enhancing the network's competence to capture long-term information within the data. Regarding SSs and solar irradiance trends, SSs with one-day and one-year increments were adopted as dominant cycles of the process, determined using scalogram and periodogram. This methodology helps in revealing and incorporating the significant cycles related to daily and annual variations in the solar irradiance.

The second approach, named WEANN, represents a hybrid model

composed of wavelet and EANN, leveraging WT as a preprocessing technique to address non-stationary and seasonal processes. Specifically, Discrete Wavelet Transform (DWT) is used to decompose the irradiance time series into subseries with various periods. The WT excels at breaking down time series into distinct sub-series, each showcasing different characteristics like frequency or seasonality. This wavelet-based decomposition simplifies the initial data, enabling the analysis of multi-frequency features within the acquired sub-series. By assigning specific weight to each sub-series, the network efficiently captures time series features, enhancing the training step and overall modeling accuracy. Wavelets, acting as translated and scaled replicas of a finite-length oscillating waveform known as the "mother wavelet," possess distinctive features like being zero mean, irregularly shaped, and compactly supported. These attributes make them well-suited for analyzing non-stationary series through WT. While the coefficients estimated by WT offer substantial information, obtaining coefficients at all scales can be resource-intensive. To address this, dyadic DWT is employed, allowing positions and scales to be chosen based on powers of 2, streamlining the analysis process. Dyadic DWT is formulated as (Addison et al., 2001):

$$\Psi_{m,p}\left(\frac{t-\tau}{s}\right) = s_0^{-\frac{m}{2}}\Psi\left(\frac{t-p\tau_0 s_0^m}{s_0^m}\right) \quad (6)$$

Where p and m are respectively translation (time) and the wavelet dilation (scale); τ_0 denotes the location parameter greater than zero; s_0 stands for specified dilation value that must be larger than 1. As an easy option, τ_0 and s_0 can be assigned by values 1 and 2, respectively. In this representation, a wavelet is composed of child wavelet sets ($\Psi_{m,p}$), obtained from the primary function Ψ, referred to the mother wavelet (Addison et al., 2001).

In the WEANN model, solar irradiance data is initially decomposed using WT into small-scale fluctuations and a large-scale trend component to capture temporal features. Subsequently, these decomposed subseries are selectively fed into the EANN model, ensuring that only dominant and informative signals are included. This approach reduces noise and redundant information, optimizing the model's structure and enhancing its performance.

The performance of both SEANN and WEANN models is compared to EANN and classic Feed Forward Neural Network (FFNN) models to evaluate the effectiveness of SS-based irradiance modeling. This comparative analysis helps to assess the potency of these models in prediction of solar irradiance patterns.

2.3. Structure of classic FFNN model

The FFNN, utilizing its Backpropagation (BP) learning algorithm, has been successfully applied to diverse engineering problems. In ANNs, "feed-forward" signifies that information flows through the network without backward connections, adept at capturing intricate data relationships for predictive modeling.

As mentioned previously, point prediction lacks information about modeling uncertainty. Therefore, in this study, the estimation of PIs was employed to explore and quantify the uncertainty in the modeling process.

2.4. Estimation of PIs

The AI techniques provide single-point predictions without accounting for uncertainties inherent in the data's stochastic nature (see Fig. 4). This limitation undermines their utility in engineering decision-making, where designs typically hinge on statistical intervals covering minimum and maximum values. Reliable uncertainty quantification in irradiance modeling is essential due to these diverse sources of uncertainty, offering crucial insights for decision-makers.

Model uncertainty arises from two main sources: modeling error and prediction error. Modeling error stems from simplifications and weaknesses in the model, while prediction error arises from inherent randomness in the data and modeling process (Morgan and Henrion, 1990). In the present study, to quantify the uncertainty associated with the modeling, the concept of PIs was applied using the common bootstrap technique (Khosravi et al., 2011). The bootstrap approach, central to this investigation, entails replicating training sets to generate multiple models. This method illuminates the variability and characteristics of an unknown distribution, providing a practical means to assess uncertainty in neural network modeling. However, traditional bootstrap methods become impractical for seasonal-based models (Singh, 1981). To address this limitation, the MBB, an enhanced version of the traditional bootstrap method, has been introduced. MBB preserves the sequence of data by selecting blocks of data in a sequential manner rather than randomly choosing individual data points for resampling (Liu and Singh, 1992; Kunsch, 1989). Consequently, MBB has demonstrated increased accuracy in time series modeling compared to the classic bootstrap, and notably, it is well-suited for application to seasonal-based models.

If the nth observation in the time series is denoted as X_n. Considering overlapping blocks ($B1,\ldots,Bn$) of length l that $1 \leq l \leq n$, each containing X_n, then:

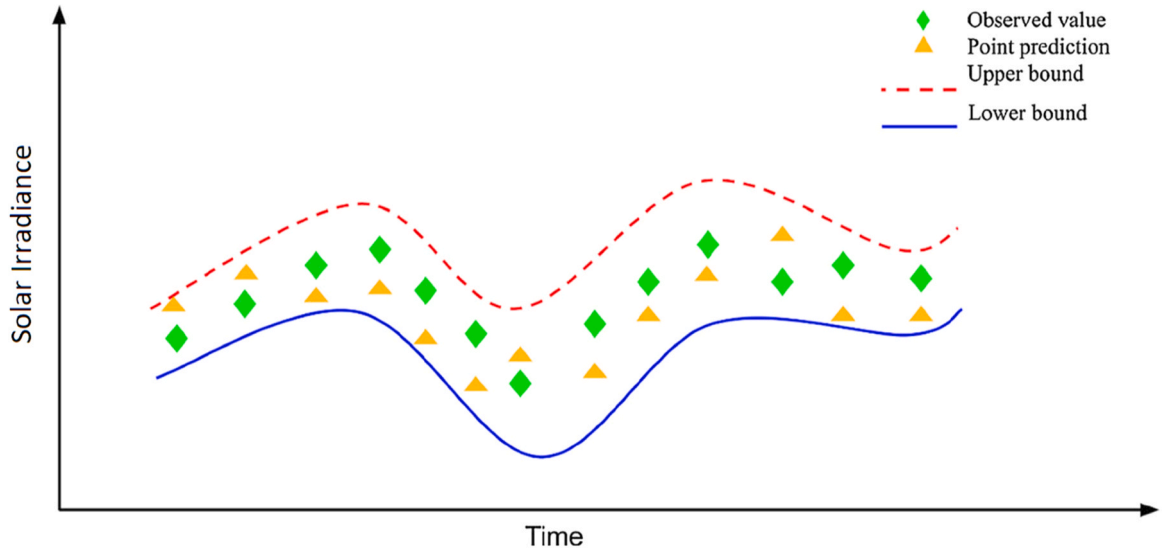


Fig. 4. The schematic comparison of PIs and point predictions for solar irradiance.

$$B1 = (X_1.X_2.\ldots.X_l), \quad B2 = (X_2.X_3.\ldots.X_{l+1}), BN = (X_{n-l+1}.\ldots.X_n) \tag{7}$$

where $N = n - l + 1$, MBB samples are generated by selecting b blocks with replacement from the set ($B1,B2,\ldots, BN$), each of size l. Therefore, during resampling b blocks of length l yield $b.l$ bootstrap observations, represented as $X^*_1.\ldots.X^*_n$. Notably, setting l to 1 transforms the MBB technique into a classic bootstrap method.

In implementing MBB for EANN-based and FFNN models, b blocks of length l are constructed, and B of such blocks are selected with replacement. A network is then trained for each of these b blocks. Collecting predictions from each trained network yields B. l sets of estimations, with l representing the length of the estimation period (see Fig. 5). Finally, statistical measures such as variance ($\hat{\sigma}^2_{boot}$) and mean ($\hat{y}_{boot}$) are computed for the estimations derived from these B sets of ANNs. A model as $f_{ANN}(x)$ is fitted to each of the generated bootstrap sub-sets, and the bootstrapping estimate is calculated as the average and variance of each model as:

$$\hat{y}_{boot}(\mathrm{x}) = \frac{1}{\mathrm{B}}\sum_{b=1}^{B} f^b_{ANN}(x) \tag{8}$$

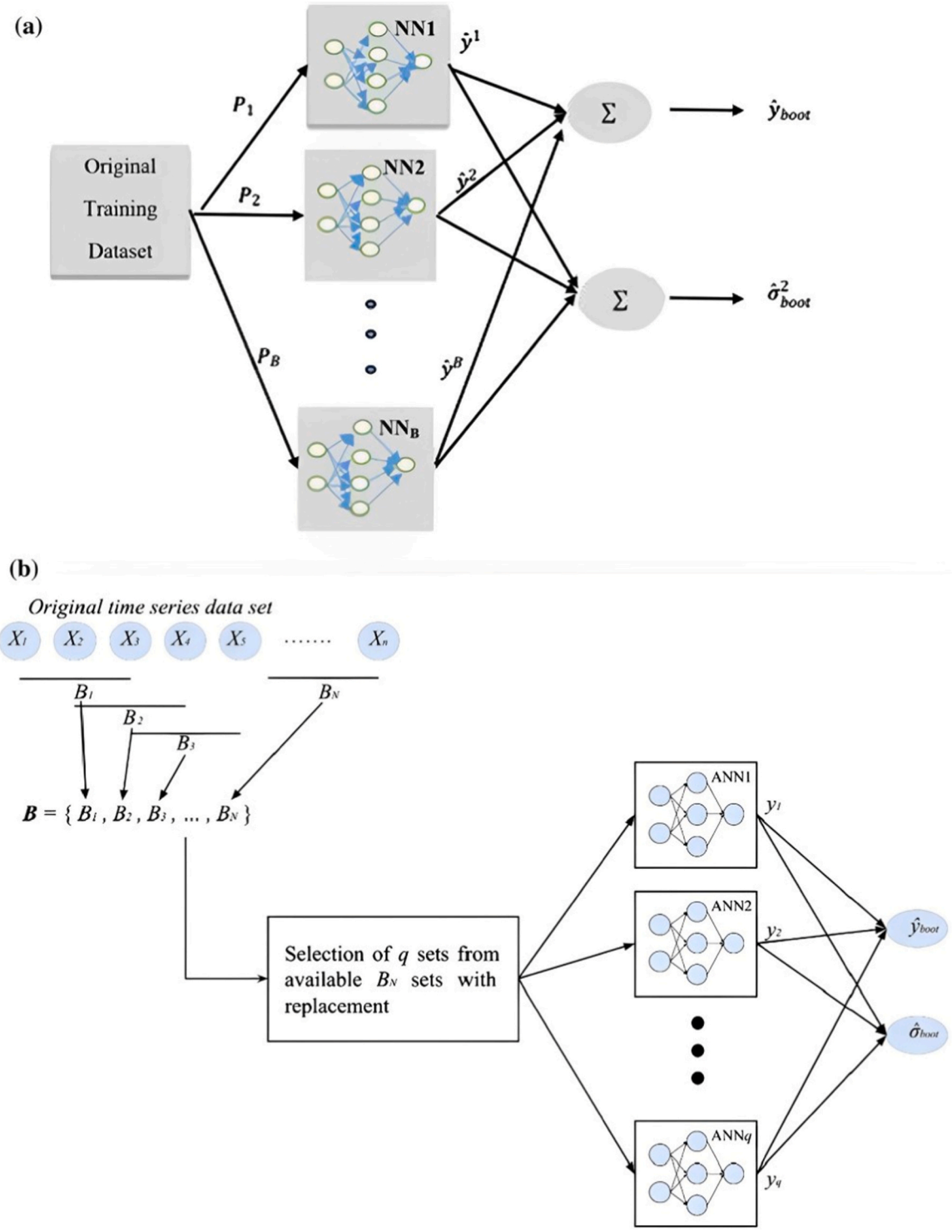


Fig. 5. The diagram of ANN-based PI method of **(a)** classic bootstrap method and **(b)** proposed MBB method.

$$\widehat{\sigma}^2_{boot}(\mathrm{x}) = \frac{1}{\mathrm{B}-1}\sum_{b=1}^{B}\left(f^b_{ANN}(x) - \widehat{y}_{boot}(x)\right)^2 \quad (9)$$

For constructing the lower and upper bands of PIs $[l.u]$, values of the recorded $X = (x_1, x_2, \ldots, x_3)$ with normal distribution probability of P and according to Eq. (10) and (11), $P\ (l < x < u)$ are as:

$$P(l < x < u) = P\left(\frac{l - \widehat{y}_{boot}}{\widehat{\sigma}_{boot}} < \frac{X - \widehat{y}_{boot}}{\widehat{\sigma}_{boot}} < \frac{u - \widehat{y}_{boot}}{\widehat{\sigma}_{boot}}\right) \quad (10)$$

where $Z = \frac{x - \widehat{y}_{boot}}{\widehat{\sigma}_{boot}}$, the standard score of x, is distributed as standard normal, hence:

$$\frac{l - \widehat{y}_{boot}}{\widehat{\sigma}_{boot}} = -Z; \frac{u - \widehat{y}_{boot}}{\widehat{\sigma}_{boot}} = Z \quad (11)$$

For further details about PIs, refer to Kreiss and Lahiri (2012) and Liu and Singh (1992).

2.5. Efficiency criteria

As a measure of the efficiency of the methods, the Determination Coefficient (DC), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were evaluated. These performance metrics are commonly used to assess the accuracy and precision of predictive models. The DC provides insight into the proportion of variance in the dependent variable that is predictable from the independent variable, while the RMSE quantifies the average magnitude of the model's errors (Nourani, 2017):

$$DC = 1 - \frac{\sum_{i=1}^{N}\left(S_{obs_i} - S_{com_i}\right)^2}{\sum_{i=1}^{n}\left(S_{obs_i} - \overline{S}_{obs}\right)^2} \quad (12)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(S_{obs_i} - S_{com_i}\right)^2}{N}} \quad (13)$$

$$MAE = \frac{\sum_{i=1}^{N}\left(\left|S_{obs_i} - S_{com_i}\right|\right)}{N} \quad (14)$$

In which N demonstrates number of data, S_{obs_i} is the observed data, and S_{com_i} denotes the output (estimated) data. DC ranges from -∞ to 1, where the value of 1 indicates a perfect modeling, and RMSE differ between 0 to +∞, where the value of 0 refers to an impeccable model.

Two key metrics for the quantitative evaluation of generated PIs are the mean PI width (MPIW) and PI coverage probability (PICP). The PICP measures the percentage of recorded values that lie within the prediction band. A wider prediction band encompasses a higher proportion of recorded values. The reliability of PIs is indicated by how closely the PICP aligns with its nominal value. The PICP is calculated using Eq. (15) and serves as a crucial measure of PI performance:

$$PICP = \frac{1}{N}\sum_{i=1}^{N} c_i \quad (15)$$

if $l(x_i) < xi < u(x_i) \rightarrow \mathrm{c_i} = 1$; else $\rightarrow \mathrm{c_i} = 0$

where $u(x_i)$ and $l(x_i)$ represent respectively the upper and lower bounds of PIs corresponding to the i th sample.

The optimal PICP value could be achieved if the upper and lower bounds of PIs had exceedingly large values. However, excessively wide PIs lack value as they provide no information regarding the recorded values. Therefore, an additional criterion is necessary to quantify the width of prediction bounds, as outlined in Eq. (16) (Khosravi et al., 2011):

$$MPIW = \frac{1}{N}\sum_{i=1}^{N}\left(l(x_i) - \quad u(x_i) \quad\right) \quad (16)$$

Subsequently, to facilitate the comparison of PI widths obtained through various techniques, it is essential to normalize the value of MPIW as follows:

$NMPIW = MPIW / R$ (17)

in this equation, R represents the range of the underlying target. Normalized MPIW (NMPIW) is a dimensionless measure that indicates the narrowness of PIs. It is evident that if NMPIW becomes excessively large, PICP will approach 1. However, as PICP alone provides no information about recorded values and overly large NMPIW is not desirable for useful estimation, both criteria (PICP and NMPIW) should be used in tandem to evaluate PIs.

3. Results and discussion

In the present study, neural network models, a conventional FFNN and two more sophisticated seasonal-based EANN, were employed to simulate solar irradiance across six stations in Iran and USA. The FFNN consisted of three layers - one input layer, one hidden layer, and one output layer - and was trained using the Levenberg-Marquardt algorithm of BP. Similarly, the solar irradiance was modeled in the second step using a simplified version of EANN with the BP method. The model was constructed based on a robust framework to ensure reliable results. A sensitivity analysis was conducted by adjusting parameters such as the percentage of data for training (60–80), activation function (tansig, logsig, purelin), training epochs (10–100), number of hidden neurons (2–30), and emotional hormones (1–20). After normalizing the data, SS sequences were generated and used as inputs to the SEANN model. Therefore, to extract the seasonal components of the radiation time series, periodogram and scalogram methods were used. The periodogram, a tool derived from Fourier analysis, is utilized to identify dominant frequencies within a time series, highlighting the strength of different frequencies present in the signal. On the other hand, the scalogram, associated with wavelet analysis, offers a time-frequency representation that enables the examination of varying frequencies over time. Fourier analysis is a transformative tool that converts signals from their primary domain to the frequency domain. Time series can be expressed using sine and cosine functions with varying periods and amplitudes, enabling the assessment of periodic attributes. A periodogram evaluates significant frequencies in time-series data, aiding in the detection of inherent periodic signals. It is a nonparametric assessment of the power spectral density of a stationary random process, often represented using Fast Fourier Transform. Fig. 6 displays the periodogram of solar irradiance for the Tabriz station, illustrating frequency components and their magnitudes.

According to Fig. 6, the normalized frequencies for the first, second, and third peaks are 0.0002289, 0.083272, and 0.16679, respectively. The corresponding periods for these peaks are 8737.44 hours (one year), 24 hours (one day), and 12 hours, respectively. It is worth noting that in Fig. 6, where the frequency unit is π rad/sec, the period time is obtained by multiplying the period values by 2, as 2π is used in the transform. The periodogram highlights extremums where the maximum and minimum values of irradiance are delineated. Subsequently, peak points in the Fig. 6 indicate high frequencies, specifically the 12-hour points that represent times with no solar irradiance at night. These 12-hour frequencies should not be considered as predominant frequencies for modeling non-zero solar irradiance. In contrast, the irradiance time series include both maximum and minimum (zero) values, generating periodicities of one day and one year. The minimum and zero values contribute to other periodicities, such as 12 hours, which are not incorporated into the modeling procedure (it is important to note that the irradiance time series assigns zero values every 12 hours). Therefore, in the modeling process, the focus should be on the maximum values as key points, while the minimum values (equal to zero at night) can be

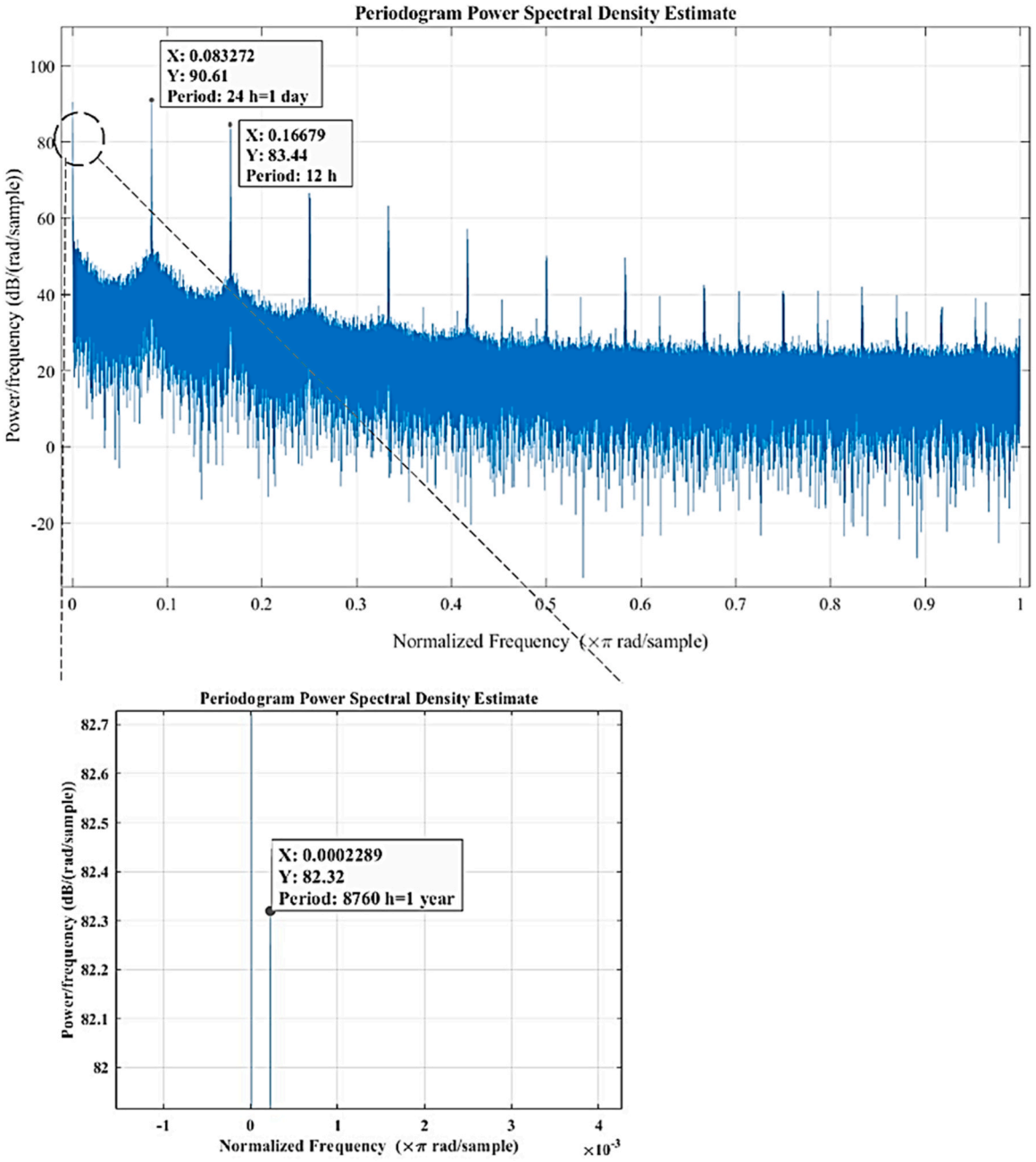


Fig. 6. Periodogram of solar irradiance time series for Tabriz station.

dismissed. This approach ensures a more accurate representation of the solar irradiance patterns.

While the periodogram provides insight into the frequency components of a time series, it may struggle with time-localized variations, whereas the scalogram excels in revealing both frequency and temporal characteristics, making it adept at capturing transient phenomena. Thus, the use of scalogram enhances the comprehensive understanding of time series data, allowing for a nuanced exploration of its frequency domain properties. WT, regarded as a breakthrough in signal processing analysis, is a practical tool for examining non-stationary oscillations through the application of mathematical transformations (Nourani et al., 2014). A Scalogram, as a time-frequency plane decomposition of time series, is achieved through CWT, illustrating the percentage of energy for each coefficient in the matrix of the CWT. CWT is an effective method for signal analysis in the time-frequency domain, providing good localization in both time and frequency domains. In this study, CWT was employed as it identifies significant characteristics of a signal that might be neglected by other methods. Thorough analysis of a signal, including the revelation of dominant frequencies is a fundamental step. Scalograms of solar irradiance are presented in Fig. 7 for six stations, aiming to investigate the potent frequencies in the solar irradiance process. The scalograms visually represent the time-frequency characteristics of the solar irradiance signals at stations, offering insights into the dominant frequencies and their temporal variations.

According to Fig. 7, it is evident that data in all six stations exhibit stationarity. Furthermore, the key periods and frequencies of the solar irradiance process are identified as 12 hours, 24 hours (a day), and 8760 hours (a year). These results align with the findings from the periodogram analysis.

The comparison of results from both the periodogram and scalogram indicates that the dominant seasonal periods in hourly solar irradiance time series are primarily associated with 1-day and 1-year periods.

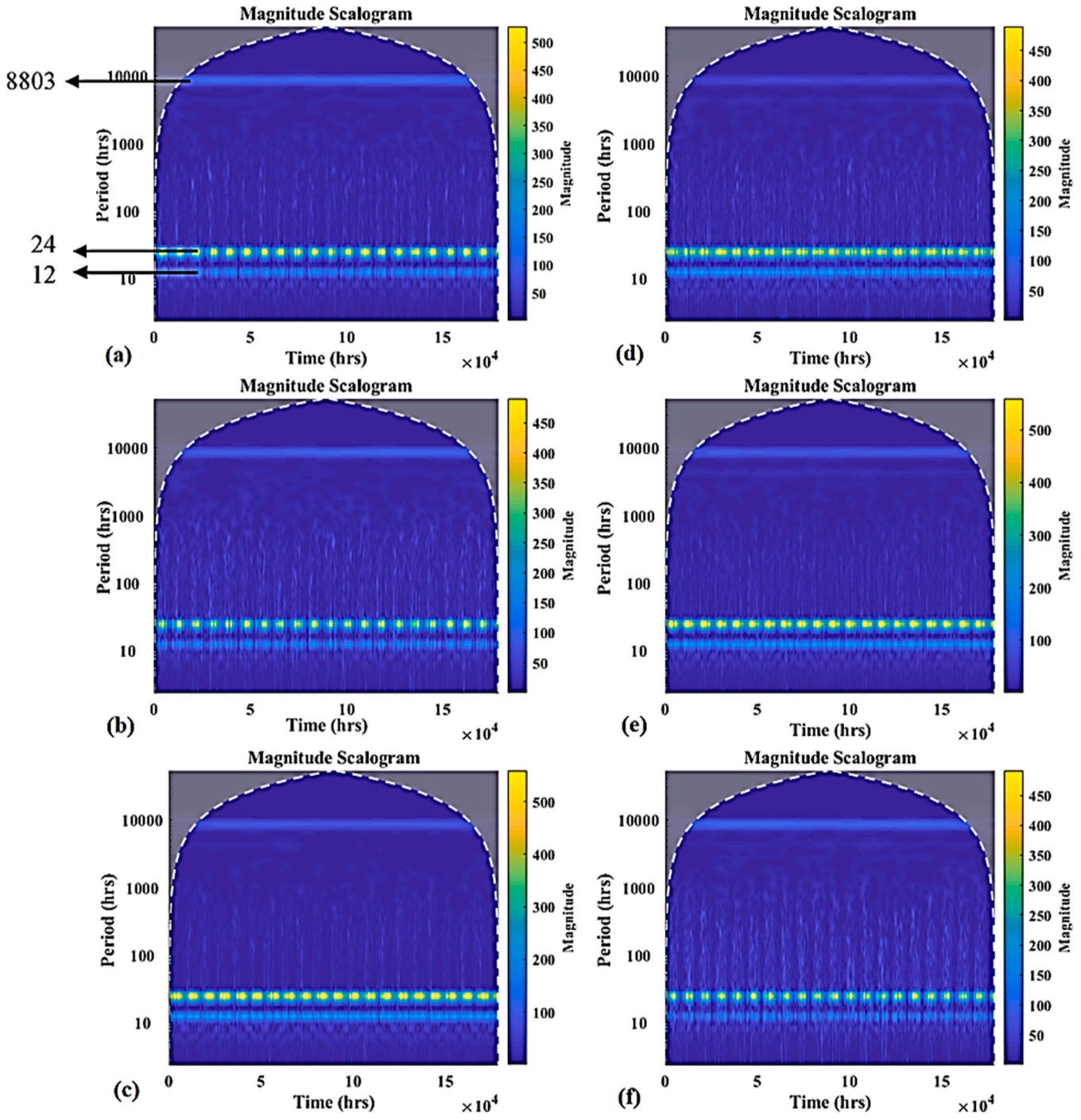


Fig. 7. Scalograms of solar irradiance for **(a)** Tabriz, **(b)** Rasht, **(c)** Zahedan, **(d)** Miami, **(e)** Tucson and **(d)** Minneapolis stations.

These periods were selected for utilization in the SEANN modeling approach. The consistency between the results obtained from the periodogram and scalogram reinforces the reliability of the identified dominant frequencies in the solar irradiance patterns and confirming the stationary of the data. Temperature and cloud cover time series were also used as auxiliary inputs to SEANN. It is noteworthy that SS was exclusively considered for the solar irradiance time series. This choice was made based on the fact that by capturing the main cycle (seasonality) of solar irradiance, the seasonality of climatic parameters is inherently accounted for as well. This approach enhances the model's ability to identify and leverage short and long-term temporal dependencies in the solar irradiance patterns.

To construct the WEANN model, the solar irradiance time series were processed at level 5 using the Meyer mother wavelet. This processing resulted in six subseries: one approximation and five detailed subseries. The selected time series were separated into trend terms and periodic terms using the WT. The sub-series obtained through WT were then designated as candidate inputs to the EANN model. The choice of the optimum decomposition level was determined using a formula where the minimum level of composition, denoted as L, is linked to the number of data points within the time series (N) (Aussem et al., 1998, Nourani et al., 2014):

$$L = int\ [log\ (N)] \tag{18}$$

Here $N = 87600$, so $L = 4$. This selection ensures an effective decomposition that captures relevant features in the solar irradiance patterns. In this study, determining the decomposition level (L) was the initial criterion for performing the WT decomposition. We explored resolution levels 2–6 using the proposed models and selected the decomposition level 5, which led to the optimal result. Several mother wavelets were examined, and the Meyer mother wavelet was identified as the most suitable for the study. It is important to note that WT was specifically applied to decompose the solar irradiance time series. The influence of seasonal qualities in climate data was indirectly considered in the modeling procedure, accounting for the frequency of solar irradiance. Additionally, normalization to the ranges (0,1) and (-1,1) was applied to the approximation and detail sub-series. The normalized solar irradiance sub-series were then designated as candidate inputs for the WEANN. Next, the wavelet energies were utilized to identify the dominant sub-series to be fed into the WEANN. The energy of each sub-series, which measures the magnitude of the wavelet coefficients in a sub-series at level t is obtained by (Kováč et al., 2021):

$$E = \sum_N |C_t|^2 \tag{19}$$

where N equals to the number of wavelets and C_t equals to the wavelet coefficients. The criteria for selecting sub-series involve identifying those with higher wavelet energies (i.e. larger wavelet coefficients), indicating stronger representation of relevant temporal features. For instance, Fig. 8 illustrates the resulting sub-series of solar irradiance data for the Tabriz station at level 5. Here, *I* represents the solar irradiance data, while *di* and *a5* correspond to the detail and approximation sub-series, respectively, with the selected sub-series marked. This approach ensures that only dominant and informative signals are included in the EANN model, minimizing noise and redundant information.

Lastly, along with temperature and cloud cover, three dominant sub-series with higher energies were chosen as inputs to the EANN model. In the final evaluation, the performance of the proposed techniques in capturing short and long-term information was assessed by comparing the seasonal-based EANN models with those of the classic FFNN model and the sole EANN.

A detailed summary of the parameters and hyperparameters used in each model is tabulated in Table 2 and the results of the proposed EANN and conventional FFNN models are summarized in Table 3. The best-performing and most accurate model is indicated in bold. The estimated and observed solar irradiance time series in the period from 15 May 2020–22 May 2020 for all methods are presented in Fig. 9. Additionally, box plots illustrating the performance of the employed models for Tabriz and Tucson stations (for examples) in the period from 1 June

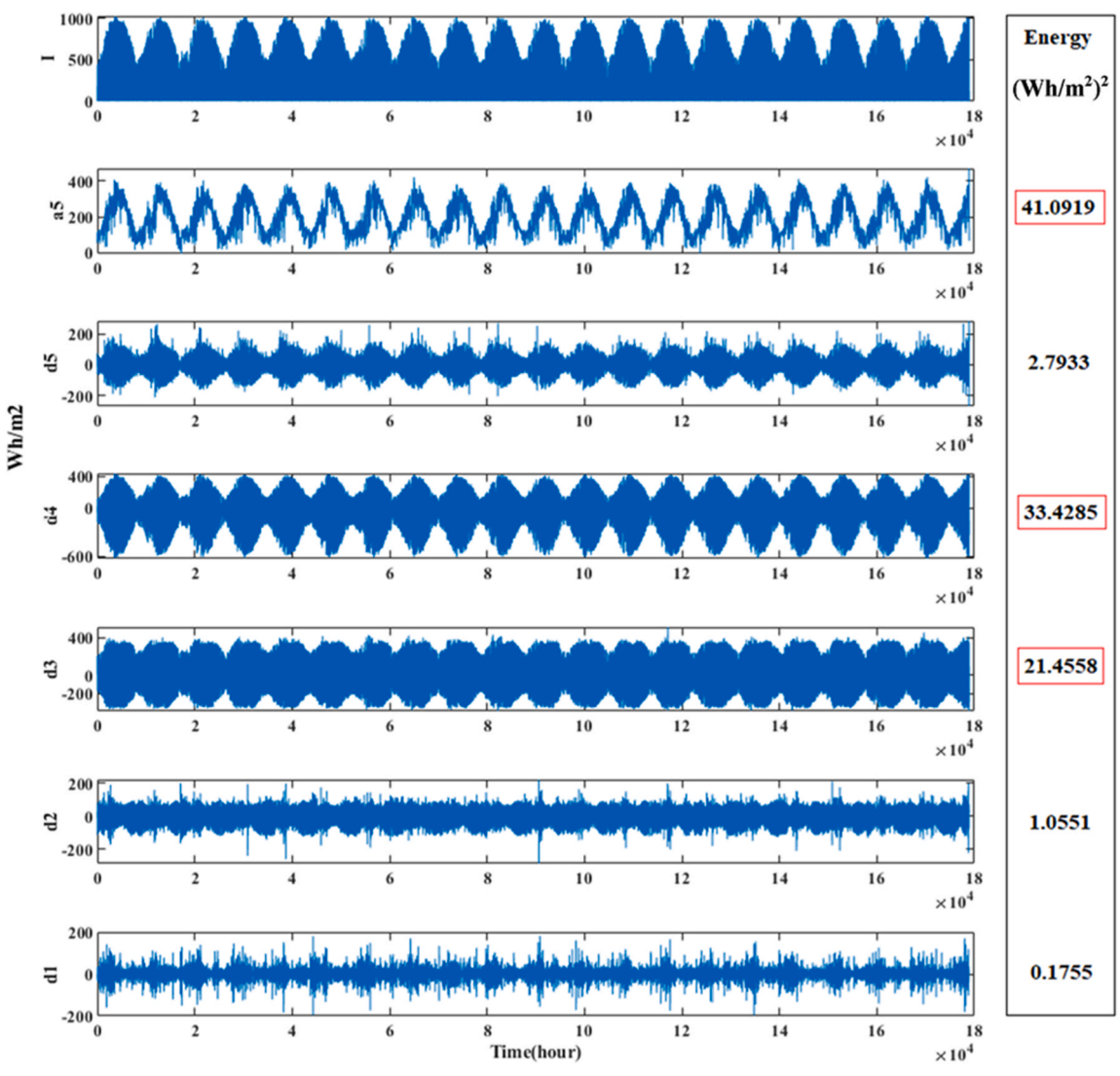


Fig. 8. Approximation and detail sub-series, along with their respective energies, for the solar irradiance time series at Tabriz station at level 5.

Table 2
Parameters and hyperparameters used in EANN-based models.

Activation Function	Hyperbolic tangent sigmoid
Learning Rate	0.01
Training Algorithm	Algorithm of BP
Number of Layers	3
Number of Neurons	Between 1 and 15 depending on the optimal calibration
Number of Hormones	Between 1 and 25 depending on the optimal calibration
Number of Epochs	Between 10 and 200 depending on the optimal calibration
Mother Wavelet (for WEANN)	Meyer

2018–30 September 2018 are shown in Fig. 10. In Fig. 10, the upper line represents the maximum, the lower line represents the minimum, and the line within the boxes represents the median. As evident from Figs. 9 and 10, while all employed models yielded acceptable outcomes, the proposed WEANN model, in particular, provided more accurate outputs. Based on the presented box plots, the results of WEANN model exhibit a distribution similar to the observed data, emphasizing the effectiveness of the proposed approach.

According to the findings presented in Table 3 and the comparison among the four stations, it can be concluded that Tucson and Zahedan produced the most accurate results, while Tabriz and Miami demonstrated performance that was more convincing than Rasht and Minneapolis. Notably, Rasht, Tabriz, and Minneapolis are located in northern latitudes, experiencing lower solar irradiance, whereas Miami, Zahedan, and Tucson are situated in lower latitudes, resulting in higher solar irradiance. Consequently, the modeling in stations with higher solar irradiance proved to be more precise and effective compared to those with a lower rate of solar irradiance. The geographical location and the presence of cloudy days make modeling for cities in northern latitudes more challenging and less reliable, as evidenced by the performance disparities observed across the stations.

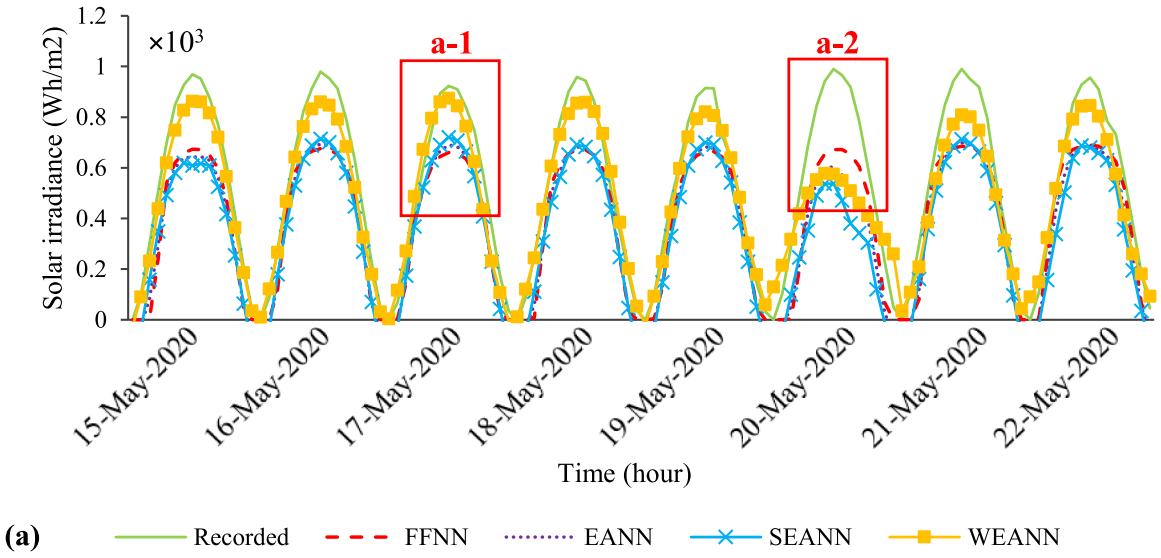
The results presented in Table 3 demonstrate that the seasonal-based EANN models outperformed both sole EANN and FFNN models. In terms of DC, the seasonal-based EANN models exhibited superior performance compared to the sole EANN outputs by up to 15 % and demonstrated greater accuracy than FFNN models by up to 25 %. The findings suggest that the seasonal-based EANN models, particularly the WEANN models, successfully capture long temporal seasonal patterns, contributing to enhanced modeling accuracy. To be more precise, the utilization of a seasonal-based EANN model has demonstrated notable enhancements in EANN modeling performance, with an increase in DC up to 9 %, 15 %, and 13 % for Tabriz, Rasht, and Zahedan, respectively, during the test phase. Similarly, for Tucson, Minneapolis, and Miami, the corresponding improvements were 11 %, 14 %, and 12 %. The discernible pattern suggests that the application of seasonal-based models tends to contribute more significantly to the modeling performance in stations with lower efficiency in solar irradiance modeling, where solar irradiance is comparatively lower than in other stations. Additionally, these methods exhibit reduced reliance on the modeling area conditions, thereby facilitating more precise predictions.

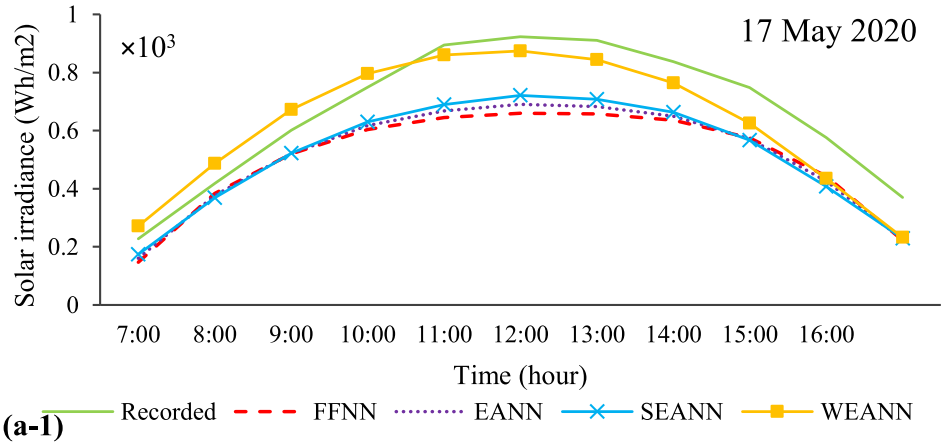
Certainly, the incorporation of emotional hormones in the EANN enhances its capability to effectively detect variations in the data, giving it an edge over the FFNN model. The effectiveness of the model used in this study is evident in its ability to describe abnormal and extreme conditions well with only a few hormonal weights, especially when combined with the WT. The WEANN model, with its external hormonal parameters and wavelet-based data pre-processing unit, demonstrates the potential to handle training iterations without requiring an excessive number of hormones. By leveraging the advantages of EANN, such as the autoregressive feature, and combining with the WT, the hybrid WEANN model overcomes limitations typically associated with classic AI models in managing non-stationary signal attributes. The superiority of WEANN over SEANN, as seen in Figs. 9 and 10, highlights the effectiveness of the seasonal models in producing more accurate results, even though almost all models led to reliable outcomes.

The analysis suggests that for cloudy days, SEANN and WEANN models exhibit results with greater precision, with a noticeable difference observed between the historical and predicted solar irradiance values for FFNN and EANN models. In contrast, the historical and predicted data values for sunny days are well-matched due to the clear and constant state of the sky. However, some discrepancies are observed in FFNN and EANN models, primarily attributed to their lower competence in evaluating peak points. Consequently, the influence of the sky state on the performance of the seasonal models proposed in this study appears

Table 3
Results of the EANN, FFNN and SEANN and WEANN models for solar irradiance modeling.

Country	Station	Model	Structure	Hormone	DC		RMSE (Wh/m^2)		MAE (Wh/m^2)		Training Time (Minute)	Memory Usage (GB)
					Train	Test	Train	Test	Train	Test		
Iran	Tabriz	FFNN	2–5–1	-	0.81	0.80	124.16	130.96	105.46	106.45	13	3.2
		EANN	2–6–1	6	0.87	0.85	90.58	98.69	79.51	79.74	31	4.9
		SEANN	4–6–1	3	0.93	0.91	75.52	82.01	63.64	65.74	73	5.6
		WEANN	**5–4–1**	**3**	**0.94**	**0.93**	**71.81**	**82.14**	**62.07**	**65.96**	**102**	**6**
	Rasht	FFNN	2–2–1	-	0.69	0.70	140.90	137.59	119.49	117.84	15	3.2
		EANN	2–6–1	14	0.75	0.74	112.35	111.78	97.55	93.32	31	4.9
		SEANN	4–6–1	10	0.85	0.84	89.27	96.85	71.67	83.23	75	5.6
		WEANN	**5–2–1**	**10**	**0.86**	**0.85**	**89.17**	**95.67**	**72.40**	**78.66**	**106**	**6**
	Zahedan	FFNN	2–5–1	-	0.80	0.79	152.61	152.40	128.27	127.28	12	3.2
		EANN	2–4–1	6	0.86	0.85	98.97	99.35	85.06	86.59	30	4.9
		SEANN	4–2–1	3	0.96	0.96	65.12	68.18	54.43	58.21	72	5.6
		WEANN	**5–4–1**	**3**	**0.97**	**0.96**	**63.40**	**65.12**	**51.95**	**55.28**	**103**	**6**
USA	Tucson	FFNN	2–4–1	-	0.81	0.82	143.61	140.25	124.85	116.40	13	3.2
		EANN	2–6–1	6	0.85	0.84	116.31	117.85	97.42	97.74	30	4.9
		SEANN	4–6–1	3	0.94	0.92	78.50	91.98	68.24	79.02	71	5.6
		WEANN	**5–2–1**	**3**	**0.96**	**0.93**	**75.75**	**93.47**	**61.47**	**77.45**	**101**	**6**
	Minneapolis	FFNN	2–5–1	-	0.74	0.68	134.17	145.99	108.51	119.67	16	3.2
		EANN	2–12–1	6	0.75	0.70	127.94	130.57	105.30	105.09	32	4.9
		SEANN	4–10–1	3	0.78	0.77	89.62	102.19	77.41	82.02	74	5.6
		WEANN	**5–12–1**	**3**	**0.81**	**0.80**	**84.06**	**97.03**	**70.56**	**79.18**	**106**	**6**
	Miami	FFNN	2–2–1	-	0.74	0.74	145.35	146.47	121.17	121.03	14	3.2
		EANN	2–7–1	22	0.81	0.81	108.91	109.99	90.18	93.26	30	4.9
		SEANN	4–6–1	10	0.92	0.91	87.15	88.87	73.14	72.45	72	5.6
		WEANN	**5–10–1**	**20**	**0.93**	**0.92**	**84.90**	**85.79**	**69.10**	**70.64**	**103**	**6**





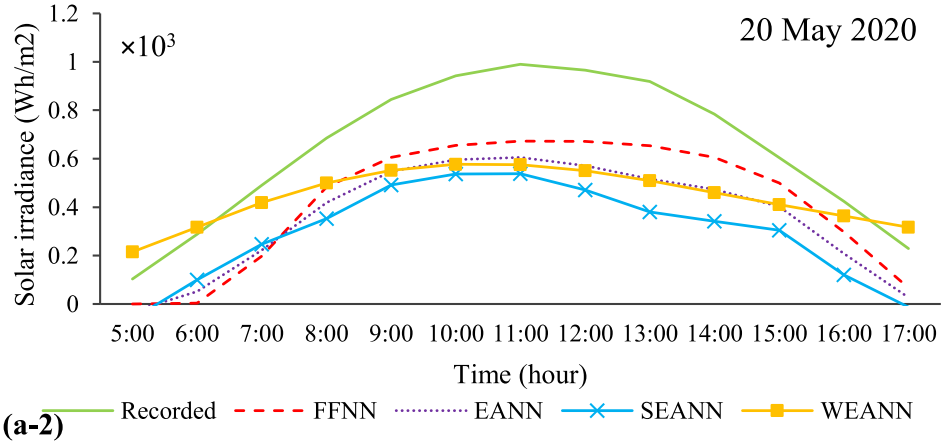
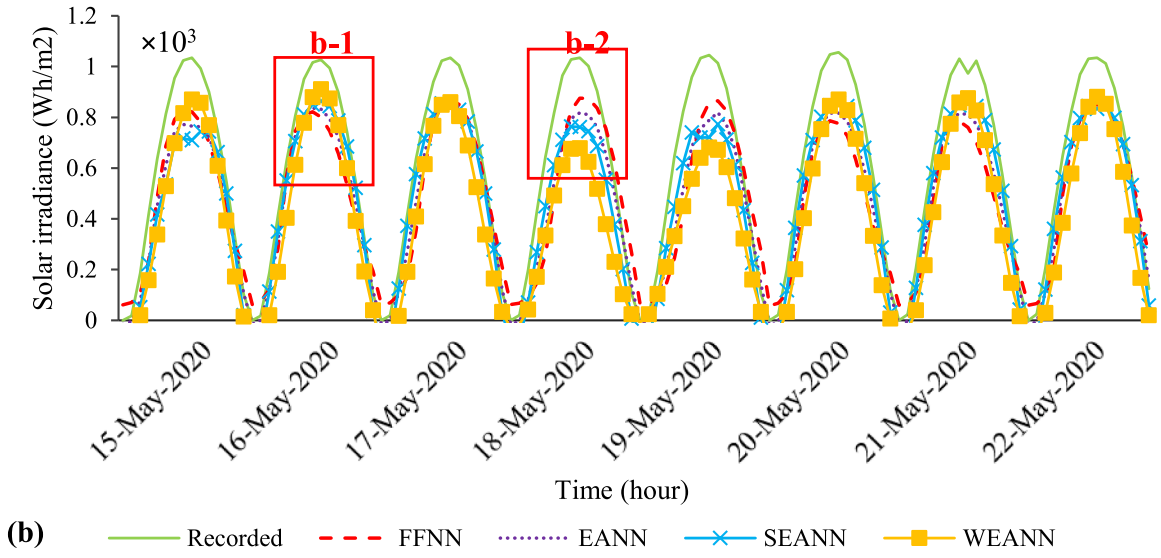
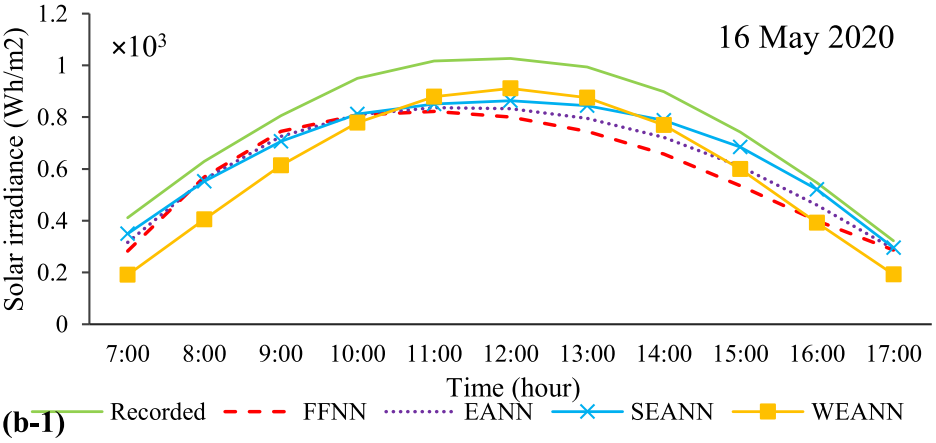


Fig. 9. Recorded and predicted values of the FFNN and seasonal EANN models for **(a)** Tabriz station and **(b)** Tucson station.

to be less pronounced.

The computational efficiency of the proposed models is an important consideration for their practical application. While the WEANN model demonstrated superior predictive performance, it also required significantly more training time and memory usage compared to the other models. This trade-off between accuracy and computational efficiency is an important factor for stakeholders to consider, especially in real-time applications or environments with limited computational resources. The SEANN model, although less computationally demanding than the WEANN model, still offered substantial improvements in predictive





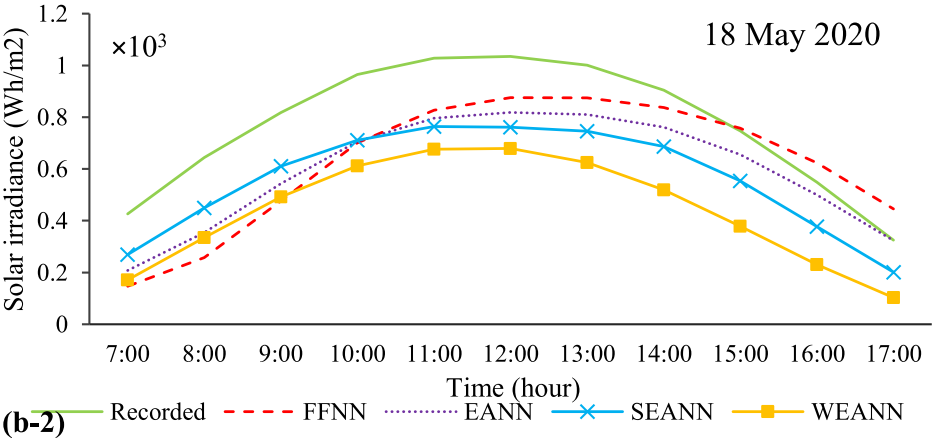


Fig. 9. (*continued*).

accuracy over the EANN and FFNN model, making it a balanced choice for many applications. The FFNN model, being the least computationally intensive, is suitable for scenarios where quick model deployment and lower resource usage are prioritized over predictive accuracy. It should be noted that this study was conducted on six stations over a period of 10 years and each station required a maximum of 2 hours for analysis, therefore, resource usage was not a significant issue. However, if similar work were to be performed on a gridded global scale and over a longer time period, advanced cloud computing and parallel computing would be necessary to manage the increased computational demands.

Also, Taylor diagrams are provided for Tabriz and Tucson stations to present a more comparative comprehension of the effectiveness and

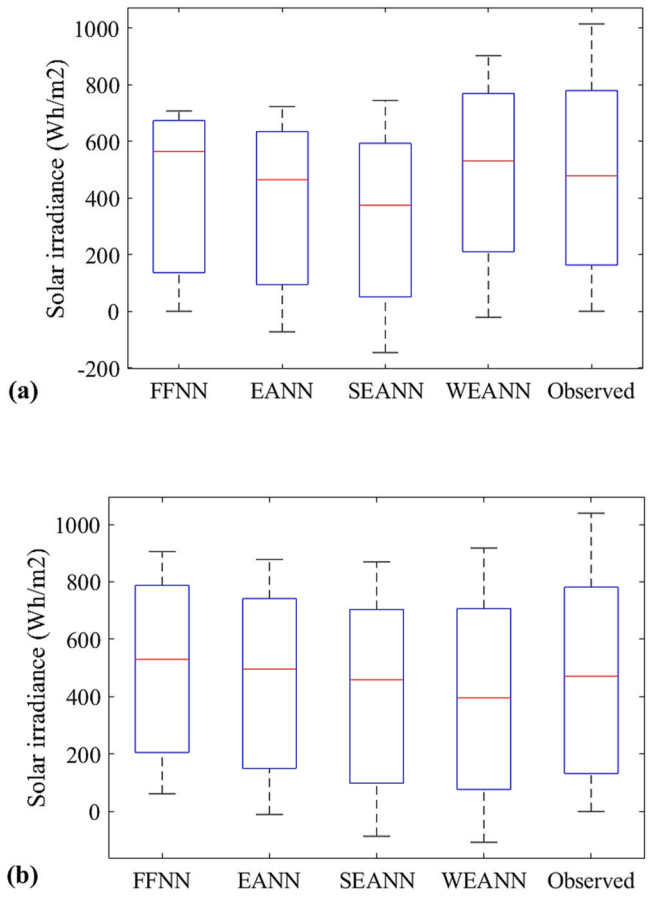


Fig. 10. Box plots of employed models for **(a)** Tabriz station and **(b)** Tucson station.

accuracy of the models in Fig. 11. Taylor diagrams typically illustrate how modeled and observed data align based on various statistical criteria, including correlation coefficients, standard deviations, and RMSE, providing a comprehensive assessment of the model performance (Taylor, 2001). Fig. 11 shows that seasonal-based EANN models could yield better predictions with higher accuracy.

In the last step and to estimate PI for uncertainty assessment of the models, the optimal hyperparameters identified for neural network models in point prediction were also applied for the calculation of PIs. The choice of confidence level in statistical methods, including ANN-based bootstrap methods, is often related to the balance between precision and reliability of results. The confidence level represents the probability that the true parameter lies within a specified interval. Higher confidence levels (e.g., 90 %, 95 %) provide wider intervals but offer greater assurance that the true parameter falls within that interval. Lower confidence levels (e.g., 80 %) result in narrower intervals but with less certainty about capturing the true parameter. For critical applications, researchers might opt for higher confidence levels to reduce the risk of making a false inference. Higher confidence levels often require larger sample sizes or more bootstrap resampling iterations, leading to increased computational requirements. In some cases, practical constraints may limit the choice of a higher confidence level. There is often a trade-off between the precision of the estimate and the width of the CI. Researchers need to consider the balance that best suits their objectives. Decision-makers might have different risk tolerances. Some may be more conservative and prefer higher confidence levels, while others may be willing to accept more uncertainty for the sake of narrower intervals. It is crucial to note that there is no one-size-fits-all answer, and the choice of confidence level depends on the specific context of the study and the goals of the analysis. In this study, confidence level for all PIs was set to 80 % which means $z = 1.282$ for MBB method to achieve narrower PIs. It was demonstrated that the number of resamples was not as critical as the sample size for MBB (Nourani et al., 2022). Thus in this study, 200 bootstrap samples were generated from the original dataset. Each bootstrap sample was used to train a separate instance of the ANN models. This moderate number strikes a balance between capturing variability and computational feasibility, and a total of 30000 samples were utilized for MBB application. In Table 4, the results of the predicted PIs with an 80 % nominal confidence level during the test step are presented. Notably, MBB computation was performed for FFNN model as a benchmark and WEANN as the model with highest efficiency due to its resource-intensive and time-consuming nature when applied to all models.

The findings presented in Table 4 highlight that, overall, EANN-

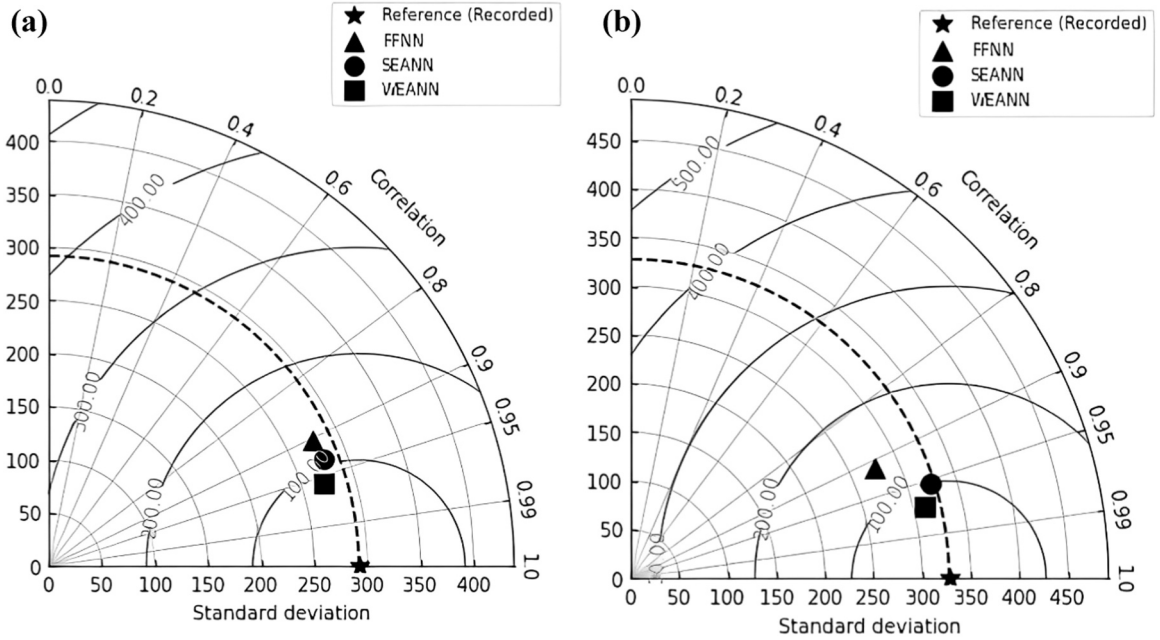


Fig. 11. Test phase Taylor diagram for **(a)** Tabriz station and **(b)** Tucson station.

Table 4
Results of estimated PIs in test step employing the autoregressive models of the EANN, FFNN and SEANN and WEANN models for solar irradiance modeling.

Station	Model	PICP	NMPIW
Tabriz	FFNN	0.633	0.910
	WEANN	**0.928**	**0.780**
Rasht	FFNN	0.168	0.961
	WEANN	**0.990**	**0.841**
Zahedan	FFNN	0.899	0.907
	WEANN	**0.951**	**0.769**
Tucson	FFNN	0.371	0.961
	WEANN	**0.914**	**0.817**
Minneapolis	FFNN	0.752	0.851
	WEANN	**0.999**	**0.721**
Miami	FFNN	0.970	0.904
	WEANN	**0.999**	**0.881**

based-MBB produced results with lower uncertainty compared to FFNN-MBB, the latter exhibiting wider PI widths. This limitation in FFNN-MBB is likely attributed to the absence of time dependency in FFNN modeling, leading to larger standard deviations (σ) and broader PIs. Conversely, seasonal-based-EANN-MBB displayed greater efficiency than EANN-MBB, considering the influence of dominant frequencies in modeling. Furthermore, WEANN-MBB outperformed all other models, exhibiting higher DC, lower RMSE and narrower width, establishing it as the most reliable model. The enhanced performance of WEANN-MBB can be attributed to the dynamic predictive capabilities of EANN, coupled with its ability to capture dominant frequencies. Another contributing factor to the superior performance of WEANN-MBB is the reliability of WEANN in providing precise point estimations, resulting in fewer unreliable networks and prediction inaccuracies during MBB. This leads to reduced σ, increased precision ŷ, and consequently narrower PIs. These characteristics result in a consistent expansion of PI width from earlier to later time steps in the test subset and a smoother overall behavior. Additionally, WEANN-MBB demonstrated the ability to capture both maximum and minimum values, rendering it a more reliable method for decision-makers.

While the integration of PIs introduced a dimension of uncertainty quantification to the proposed framework, it is important to discuss the limitations of the PIs used in this study. One limitation is that the accuracy of PIs depends heavily on the quality and quantity of the training data. In regions with limited historical solar irradiance data, the PIs may not be as reliable. Choosing an appropriate block length is crucial; too short blocks may not capture the data's autocorrelation structure adequately, while too long blocks can introduce bias. Additionally, the bootstrap method used for estimating PIs assumes that the training data adequately represents the variability in future data, which may not always be the case, especially in areas with highly unpredictable weather patterns. Moreover, the PIs might not capture extreme events or sudden changes in weather conditions effectively. For instance, in regions with frequent and abrupt changes in cloud cover, the PIs might underestimate the actual variability, leading to less reliable predictions. These scenarios highlight the need for further research to enhance the robustness of PIs, possibly by integrating more sophisticated methods or additional climatic variables. Overall, while the PIs provided by the WEANN model offer a robust measure of prediction uncertainty in many cases, their reliability can be compromised in scenarios with insufficient or highly variable data. Furthermore, MBB's computational demands can be high, especially with large datasets. Lastly, MBB's performance can be sensitive to underlying model assumptions, potentially affecting the reliability of bootstrap estimates. Future research could focus on addressing these limitations by exploring alternative methods for PI estimation and validating the model in a wider range of geographical and climatic conditions.

Overall, our seasonal EANN models offer significant improvements in the practicality of solar irradiance forecasting, particularly in real-world applications. By integrating an emotional component inspired by biological systems, these models enhance adaptability and accuracy in predicting solar irradiance under varying climatic conditions. One key advantage of EANN models is their ability to capture complex seasonal variations and dependencies within solar irradiance data. Unlike traditional neural networks, EANNs adjust activation thresholds and synaptic weights based on emotional states simulated within the network, thereby optimizing performance across different seasons and geographical locations. This dynamic adaptation improves the reliability of solar irradiance predictions, crucial for applications ranging from solar energy planning to grid management and renewable energy integration. However, it is essential to consider certain limitations. EANN models require sufficient computational resources for training and deployment. Variations in satellite coverage and weather conditions can also impact model accuracy, necessitating ongoing calibration and validation against ground-truth measurements. In real-world scenarios, the practical application of EANN models depends on addressing these challenges and integrating them into existing forecasting frameworks. Future research could explore further enhancements to model robustness and scalability, ensuring their effective deployment across diverse geographical and climatic regions. By discussing both the potential applications and limitations of our EANN models, we aim to provide a comprehensive perspective on their utility in advancing solar irradiance forecasting capabilities and supporting sustainable energy solutions.

4. Conclusions

The integration of artificial emotions into ANNs has gained significant attention, driven by the crucial role that natural emotions play in the behavioral learning processes observed in animals. This study introduced a novel AI-based model in the energy field, the seasonal-EANN, for modeling the solar irradiance of six stations in two countries with distinct climate characteristics. The results indicated that the seasonal EANN models, particularly the WEANN, effectively capture intricate patterns of solar irradiance, resulting in a 15 % improvement over the classic EANN model and a 25 % improvement over the traditional FFNN model. This superiority is attributed to the capability of WT to execute multi-resolution analysis, decomposing the original time series into multi-scale sub-series, each representing a separate seasonal scale. This allows for the consideration of the multi-seasonality of the time series. The study highlights that the simulation is more accurate for areas with higher solar irradiance, while less accurate results were observed in areas with cloudy days or northern regions with lower solar irradiance. Additionally, the integration of PIs introduced a dimension of uncertainty quantification to the proposed framework. Utilizing the MMB method with ANNs allowed for robust statistical estimation of PIs. The WEANN model not only enhanced accuracy in point predictions but also improved uncertainty quantification, with an average NMPIW of 0.80 and an average PICP of 0.96. The incorporation of PIs elevates the practicality of our model, providing decision-makers with a more dependable tool to evaluate the potential range of solar irradiance values and associated uncertainties. The proposed framework benefits the renewable energy sector by providing accurate solar irradiance forecasts. Policymakers can use these for effective energy policies, while utility companies can improve grid management and reduce costs. Enhanced accuracy and uncertainty quantification aid in planning, resource allocation, and meeting energy targets, optimizing solar energy integration and ensuring grid stability. However, PI reliability can be compromised in regions with limited data or highly variable weather. These limitations underscore the need for further research to enhance the robustness and applicability of the model. Future researches could focus on validating the model in diverse climates, integrating additional climatic variables, and using General Circulation Models (GCMs) to assess climate change impacts. The development of advanced deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), could capture more complex patterns in solar irradiance data. Investigating the impact of transfer learning

could improve model performance in regions with limited historical solar irradiance data. For future studies, it should be noted that conducting similar research on a gridded global scale and over extended periods would necessitate advanced cloud computing and parallel processing to handle the increased computational demands. Exploring these technologies would be essential to manage larger datasets and more complex models efficiently.

Author statement

We declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We understand that the Corresponding Author is the sole contact for the Editorial process. She is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs.

CRediT authorship contribution statement

Vahid Nourani: Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Anne Ng:** Writing – review & editing, Validation, Data curation. **Chunwei Zhang:** Writing – review & editing, Project administration, Conceptualization. **Fahreddin Sadikoglu:** Writing – review & editing, Resources, Project administration, Funding acquisition. **Nazanin Behfar:** Writing – original draft, Methodology, Formal analysis, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

References

Abdallah, M., Mohammadi, B., Nasiri, H., Katipoğlu, O.M., Abdalla, M.A.A., Ebadzadeh, M.M., 2023. Daily global solar radiation time series prediction using variational mode decomposition combined with multi-functional recurrent fuzzy neural network and quantile regression forests algorithm. Energy Rep. 10, 4198–4217.

Abdi, J., Moshiri, B., Abdulhai, B., Sedigh, A.K., 2012. Forecasting of short-term traffic-flow based on improved neurofuzzy models via emotional temporal difference learning algorithm. Eng. Appl. Artif. Intell. 25, 1022–1042.

Addison, P.S., Murray, K.B., Watson, J.N., 2001. Wavelet transform analysis of open channel wake flows. J. Eng. Mech. 127, 58–70.

Akarslan, E., Hocaoglu, F.O., Edizkan, R., 2018. Novel short term solar irradiance forecasting models. Renew. Energy 123, 58–66.

Aussem, A., 1998. Wavelet based feature extraction and decomposition strategies for financial forecasting. Int. J. Comput. Intell. Financ. 6, 5–12.

Babaie, T., Karimizandi, R., Lucas, C., 2008. Learning based brain emotional intelligence as a new aspect for development of an alarm system. Soft Comput. 12, 857–873.

Chen, C.R., Kartini, U.T., 2017. K-nearest neighbor neural network models for very short-term global solar irradiance forecasting based on meteorological data. Energies 10, 186.

Dabuechies, I., 1990. The wavelet transform, time-frequency localization and signal analysis. IEEE Trans. Inf. Theory 36, 961–1005.

Dewangan, C.L., Singh, S.N., Chakrabarti, S., 2017, November. Solar irradiance forecasting using wavelet neural network. In 2017 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC) (pp. 1-6). San Diego, CA, USA.

Ehteram, M., Shabanian, H., 2023. Unveiling the SALSTM-M5T model and its python implementation for precise solar radiation prediction. Energy Rep. 10, 3402–3417.

Geetha, A., Santhakumar, J., Sundaram, K.M., Usha, S., Thentral, T.T., Boopathi, C.S., Ramya, R., Sathyamurthy, R., 2022. Prediction of hourly solar radiation in Tamil Nadu using ANN model with different learning algorithms. Energy Rep. 8, 664–671.

Gutierrez-Corea, F.V., Manso-Callejo, M.A., Moreno-Regidor, M.P., Manrique-Sancho, M. T., 2016. Forecasting short-term solar irradiance based on artificial neural networks and data from neighboring meteorological stations. Sol. Energy 134, 119–131.

Khaled, F., Farouk, C., Abdalah, K., Belgacem, B., Nacer, H., 2022. Wavelet Packet-Gaussian process regression multivariate and univariate model for forecasting daily solar radiation. Int. J. Energy a Clean. Environ. 23, 77–96.

Khashman, A., 2008. A modified backpropagation learning algorithm with added emotional coefficients. IEEE Trans. Neural Netw. 19, 1896–1909.

Khosravi, A., Nahavandi, S., Creighton, D., Atiya, A., 2011. Comprehensive review of neural network-based prediction intervals and new advances. IEEE Trans. Neural Netw. 22, 1341–1356.

Koo, Y., Oh, M., Kim, S.M., Park, H.D., 2020. Estimation and mapping of solar irradiance for korea by using COMS MI satellite images and an artificial neural network model. Energies 13, 301.

Kováč, S., Michaľčonok, G., Halenár, I., Važan, P., 2021. Comparison of heat demand prediction using wavelet analysis and neural network for a district heating network. Energies 14, 1545.

Kreiss, J.P., Lahiri, S.N., 2012. 1-bootstrap methods for time series. In: Subba Rao T., Subba Rao S., Rao C.R. (eds.) Handbook of Statistics, 30, 3–26. Elsevier.

Kunsch, H.R., 1989. The jackknife and the bootstrap for general stationary observations. Ann. Stat. 17, 1217–1241.

Liu, R.Y., Singh, K., 1992. Moving blocks jackknife and bootstrap capture weak dependence. Explor. Limits Bootstrap 225, 248.

Lotfi, E., Akbarzadeh-T, M.R., 2016. A winner-take-all approach to emotional neural networks with universal approximation property. Inf. Sci. 346, 369–388.

Lyu, L., Kantardzic, M., Arabmakki, E., 2014, December. Solar irradiance forecasting by using wavelet based denoising. In 2014 IEEE symposium on computational intelligence for engineering solutions (CIES) (pp. 110-116). Orlando, FL, USA.

Meenal, R., Selvakumar, A.I., 2018. Assessment of SVM, empirical and ANN based solar radiation prediction models with most influencing input parameters. Renew. Energy 121, 324–343.

Mirparizi, M., Shakeriaski, F., Salehi, F., Zhang, C., 2023. Available challenges and recent progress in carbon dioxide capture, and reusing methods toward renewable energy. Sustain. Energy Technol. Assess. 58, 103365.

Morgan, M., Henrion, M., 1990. Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge university press.

Nourani, V., 2017. An emotional ANN (EANN) approach to modeling rainfall-runoff process. J. Hydrol. 544, 267–277.

Nourani, V., Baghanam, A.H., Adamowski, J., Kisi, O., 2014. Applications of hybrid wavelet–artificial intelligence models in hydrology: a review. J. Hydrol. 514, 358–377.

Nourani, V., Elkiran, G., Abdullahi, J., Tahsin, A., 2019a. Multi-region modeling of daily global solar radiation with artificial intelligence ensemble. Nat. Resour. Res. 28, 1217–1238.

Nourani, V., Sharghi, E., Behfar, N., Zhang, Y., 2022. Multi-step-ahead solar irradiance modeling employing multi-frequency deep learning models and climatic data. Appl. Energy 315, 119069.

Pavlos, G.P., Iliopoulos, A.C., Tsoutsouras, V.G., Sarafopoulos, D.V., Sfiris, D.S., Karakatsanis, L.P., Pavlos, E.G., 2011. First and second order non-equilibrium phase transition and evidence for non-extensive Tsallis statistics in Earth's magnetosphere. Phys. A: Stat. Mech. Appl. 390, 2819–2839.

Pfister, G., McKenzie, R.L., Liley, J.B., Thomas, A., Forgan, B.W., Long, C.N., 2003. Cloud coverage based on all-sky imaging and its impact on surface solar irradiance. J. Appl. Meteorol. Climatol. 42, 1421–1434.

Piri, J., Kisi, O., 2015. Modelling solar radiation reached to the earth using ANFIS, NN-ARX, and empirical models (Case studies: Zahedan and Bojnurd stations). J. Atmos. Sol. -Terr. Phys. 123, 39–47.

Reddy, B.S.N., Pramada, S.K., Roshni, T., 2021. Monthly surface runoff prediction using artificial intelligence: a study from a tropical climate river basin. J. Earth Syst. Sci. 130, 1–15.

Sharifi, S.S., Rezaverdinejad, V., Nourani, V., 2016. Estimation of daily global solar radiation using wavelet regression, ANN, GEP and empirical models: A comparative study of selected temperature-based approaches. J. Atmos. Sol. -Terr. Phys. 149, 131–145.

Sharifi, S.S., Rezaverdinejad, V., Nourani, V., Behmanesh, J., 2022. Multi-time-step ahead daily global solar radiation forecasting: performance evaluation of wavelet-based artificial neural network model. Meteorol. Atmos. Phys. 134, 50.

Sharma, V., Yang, D., Walsh, W., Reindl, T., 2016. Short term solar irradiance forecasting using a mixed wavelet neural network. Renew. Energy 90, 481–492.

Singh, K., 1981. On the asymptotic accuracy of Efron's bootstrap. Ann. Stat. 9, 1187–1195.

Tao, H., Ewees, A.A., Al-Sulttani, A.O., Beyaztas, U., Hameed, M.M., Salih, S.Q., Armanuos, A.M., Al-Ansari, N., Voyant, C., Shahid, S., Yaseen, Z.M., 2021. Global solar radiation prediction over North Dakota using air temperature: development of novel hybrid intelligence model. Energy Rep. 7, 136–157.

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res.: Atmosph. 106, 7183–7192.

Tercha, W., Tadjer, S.A., Chekired, F., Canale, L., 2024. Machine learning-based forecasting of temperature and solar irradiance for photovoltaic systems. Energies 17, 1124.

Thenius, R., Zahadat, P., Schmickl, T., 2013, September. EMANN - a model of emotions in an artificial neural network. In ECAL 2013, twelfth European Conference, Artificial Immune, Neural and Endocrine Systems (pp. 830-837). Cambridge, England.

Torregrossa, D., Le Boudec, J.Y., Paolone, M., 2016. Model-free computation of ultra-short-term prediction intervals of solar irradiance. Sol. Energy 124, 57–67.

Wang, L., Lu, Y., Wang, Z., Li, H., Zhang, M., 2024. Hourly solar radiation estimation and uncertainty quantification using hybrid models. Renew. Sustain. Energy Rev. 202, 114727.
Yadav, A.K., Chandel, S.S., 2014. Solar radiation prediction using Artificial Neural Network techniques: A review. Renew. Sustain. Energy Rev. 33, 772–781.
Yang, H., Kurtz, B., Nguyen, D., Urquhart, B., Chow, C.W., Ghonima, M., Kleissl, J., 2014. Solar irradiance forecasting using a ground-based sky imager developed at UC San Diego. Sol. Energy 103, 502–524.
Zhang, P., Takano, H., Murata, J., 2011, September. Daily solar radiation prediction based on wavelet analysis. In SICE Annual Conference 2011 (pp. 712-717). Tokyo, Japan.