Multi-timescale Forecast of Solar Irradiance Based on Multi-task Learning and Echo State Network Approaches

Zhou Wu ⓘ, *Member, IEEE*, Qian Li, and Xiaohua Xia ⓘ, *Fellow, IEEE*

Abstract—**Solar irradiance forecast is closely related with efficiency and reliability of renewable energy systems. Multi-timescale irradiance forecast is a new and efficient way to simultaneously predict solar energy generation on different timescales for hierarchical decision making. This article newly adopts the multi-task learning mechanism to study the multi-timescale forecast for improving accuracy and computational efficiency. A novel multi-timescale (MTS) prediction framework is presented to fulfill the multi-task application, and echo state network (ESN) is studied in the proposed MTS framework. The multi-timescale ESN (MTS-ESN) is proposed to enhance the information sharing among correlated tasks. Simulation results of hourly solar data demonstrate that the proposed MTS-ESN could achieve promising performance at both hourly and daily level in parallel. The MTS-ESN outperforms the single-timescale ESN (STS-ESN), which indicates the information sharing in the multi-task learning is effective in this application.**

Index Terms—**Echo state neural network, multi-tasking model, renewable energy, smart grid, solar irradiance forecasting.**

I. INTRODUCTION

FOR GREEN and sustainable energy production, renewable energy systems have been developed rapidly and widely over the world. Amongst renewable sources, solar energy has gained the most attention, as it is easily accessible, and environmental friendly. In order to efficiently control and manage solar energy systems, the solar irradiance in the target location must be predicted in advance, and the prediction accuracy will influence the power quality, grid security, and cost.

As reviewed in [1] and [2], different techniques have been successfully developed and applied to solar irradiance forecast, including statistical methods [3], [4], artificial neural networks (e.g., backpropagation and Elman) [5], [6], machine learning algorithms [7], [8], spatial-temporal models [9], [10], data mining [11], [12], and so forth. In addition, some advanced techniques based on wavelet analysis [13], fuzzy logic [14], [15], optimization algorithms [16], and K-means approach [17] can also be hybridized with existing forecast methods for further performance improvement.

The aforementioned forecast models mainly focus on a single-timescale prediction, which means that each model (e.g., short-term, medium-term, or long-term) can only predict the future irradiance at a specific timescale (e.g., an hour, a day, and a year). We call these models as single-task models. Although single-task models could be easily and effectively implemented, they confront two challenges in certain complicated applications with hierarchical prediction tasks, e.g., smart energy hubs. Firstly, a single-task model with good performance on a certain timescale cannot be directly applied to the prediction task on other timescales, so the cumbersome way is to design many prediction models for the situation of multi-task prediction. Secondly, the performance of a single-task model on a certain timescale has a limit, when only considering the information in the single task. In other words, other correlated tasks may include useful information, which has been neglected in the single-task models. For example, the value of solar irradiance in the next hour is not only affected by the values from previous hours, but also by the values from previous days or weeks.

To handle aforementioned challenges, we propose a multi-task learning perspective for solar irradiance prediction. Multi-task learning offers the opportunity to improve the performance of one learning task, by sharing the information and knowledge learned from other correlated tasks [18]. With the perspective of multi-task learning, multi-timescale (MTS) prediction is studied in this article to predict solar irradiance on different timescales at any time instant. On the one hand, only one multitask model is required to report irradiance values of future hours, days, or weeks at a given time instant for hierarchical decision making. To set up a multi-task model instead of many single-task models could significantly enhance the computational efficiency. On the other hand, the completeness of one task is helpful to fulfill other tasks in the multi-task models, since each correlated task shares useful information. With respect to each task, the

multi-tasking model could have more accurate results than the single-task model. Therefore, the decision makers can predict solar irradiance at different timescales in parallel, and make comprehensive decisions, e.g., short-term control, middle-term scheduling, and long-term investment.

The multi-timescale echo state network (MTS-ESN) algorithm is proposed with the multi-tasking perspective, in which ESN is an effective and simple form of recurrent neural network (RNN) [19]. ESN is a kind of neural network with three layers, in which the hidden layer (called dynamical reservoir) contains a large number of sparsely connected neurons. The dynamical reservoir could encode the temporal information of input signals from low dimensional input space to high dimensional state space. Therefore, ESN has stronger expressive and mapping capabilities than traditional RNN, whose hidden layer size is in the level of tens. In addition, only the connection weights from reservoir to output layer in ESN need to be updated by linear regression algorithms. It avoids the heavy computational burden, slow convergence and local optimization of classical RNN based on gradient descent training algorithm. Due to the simple training procedure and strong expressive and mapping capabilities, ESN has been widely applied to many practical fields, including pattern recognition [20], [21], optimal energy management [22], and especially nonlinear time series prediction [23]–[25]. The proposed MTS-ESN has employed multiple reservoirs to express input dynamics on multiple timescales, and the reservoir states will be further integrated for information sharing. Compared with ESN with the single task, MTS-ESN is expected to improve the prediction accuracy and computation efficiency due to the information sharing.

In summary, the main contributions of this article are in three folds.

1) Multi-task prediction of solar irradiance is newly modeled. The correlations between each prediction task is unveiled in multi-task models, which are neglected in the single-task models.
2) A novel MTS prediction framework is presented in this article. The model can fulfill multi-task forecast at different timescales in parallel for a given time instant.
3) The proposed MTS-ESN is an effective algorithm to predict the solar irradiance of future hours, days, and other timescales.

The rest of this article is organized as follows. Section II provides a comprehensive review of irradiance forecast methods. Section III describes the multi-task perspective of solar irradiance. In Section IV, the proposed MTS-ESN algorithm is developed. Section V presents the experimental design and results. Finally, Section VI concludes this article.

II. Overview of Solar Irradiance Forecast

As reported in [1], [2], a number of solar irradiance forecast approaches have been developed, mainly including the following three kinds of models with respect to regression technique:

1) Classical statistical models: Examples include persistence (PSS) model, autoregressive moving average (ARMA) model and autoregressive integrated moving average (ARIMA) [3], [4]. For the PSS model, the predicted solar irradiance of the next time $\hat{G}_{i+1}$ is the same as the current one G_i, namely, $\hat{G}_{i+1} = G_i$. The PSS model can be taken as a benchmark for very short term forecast horizon (less than 1 h). Recently, the autoregressive with exogenous input model (ARX) is proposed as follows:

$$\begin{aligned}
\hat{G}_i = {} & \alpha_1 G_{i-1} + \alpha_2 G_{i-2} + \cdots + \alpha_p G_{i-p} \\
& + \beta_{11}\mu_1(i-d_1) + \beta_{21}\mu_1(i-d_1-1) + \cdots \\
& + \beta_{m1}\mu_1(i-d_1-m_1+1) + \cdots + \beta_{1j}\mu_j(i-d_j) \\
& + \beta_{2j}\mu_j(i-d_j-1) + \cdots \\
& + \beta_{mj}\mu_j(i-d_j-m_j+1) + \varepsilon_i
\end{aligned} \tag{1}$$

where α, β are model coefficients, ε_i is a Gaussian white noise series with zero; μ_j is related exogenous input signal, such as temperature, humidity, and so forth; j is the dimension of exogenous input, m is the order of historical exogenous input, d is the time delay between the input μ_j and output G_i; The time index i could denote a certain time interval, e.g., hourly (h), daily (d), and so forth, and G_i is the solar irradiance at the ith interval.

2) Artificial neural network (ANN) models: These ANN methods can be categorized into univariate and multivariate models. The univariate ANN models are built only based on the historical solar irradiance data [26], [27], whose mathematical expression is given as:

$$\hat{G}_{i+k} = \tilde{f}(G_i, G_{i-1}, \ldots, G_{i-m}) \tag{2}$$

where k denotes the prediction step, G_i could be the solar irradiance from a certain time interval, $\tilde{f}$ is the nonlinear mapping relationship between historical and future signals using ANN. In addition to the historical solar irradiance data, some related meteorological parameters, such as temperature, humidity, wind speed and so on, are also considered in the multivariate ANN models [28], [29]. Mathematically, let μ denote meteorological parameters, the multivariate approach can be expressed as:

$$\hat{G}_{i+k} = \tilde{f}(G_i, G_{i-1}, \ldots, G_{i-m}, \mu). \tag{3}$$

Both the univariate and multivariate models can be used for multi-step ahead prediction. In particular, considering $k = 1$, (2) and (3) are one-step forecasting models. In this article, we mainly study the artificial ANN models, in which only historical irradiance data is utilized.

3) Other hybrid models: In order to further improve the forecast performance, some advanced methodologies can be hybridized with existing models, such as data analysis, optimization algorithms, and so forth.

Aforementioned forecast models are mostly single-task models. In other words, the hourly forecast models are only responsible to report hour-ahead values, and the daily forecast models are only responsible to report day-ahead values. Such single-task models have limited performance improvement, as some useful information among correlated tasks is neglected. For example, the value of solar irradiance in the next hour is not only affected

by the values from previous hourly timescale, but also by the values from previous daily or weekly timescale.

Multi-task learning offers a novel opportunity for further performance improvement of one task, by learning multiple correlated tasks simultaneously and sharing the knowledge or information of each task [18]. The completeness of one task is helpful to fulfill other tasks in the multi-task models, since each correlated task shares useful information or knowledge. In recent years, the multi-task learning has been applied to identification and classification [30]–[32], time series prediction [33]–[35], and so forth. To our best knowledge, few results exist in solar irradiance or PV outputs prediction based multi-task learning. [36] proposes novel iterative multi-task learning for PV output a prediction, where their task is to learn the missing observations by information sharing among multiple similar solar panels.

In this article, a novel MTS-ESN model is proposed for solar irradiance prediction, based on multi-task learning mechanism. Different from [26]–[29], the proposed MTS-ESN can learn multiple tasks in parallel, making multi-task decisions in practical situations. Unlike [36], MTS-ESN considers the fact that the solar irradiance on multiple timescales are correlated. The correlations on different timescales are therefore utilized for multi-task modeling in MTS-ESN. In addition, only the historical solar irradiance data of target station is used for MTS-ESN modeling.

III. NEW PERSPECTIVE OF MULTITASK LEARNING

With the perspective of multi-task learning, the completeness of one task is helpful to fulfill other correlated tasks, since each task in the multi-task model interacts with information sharing. For these purposes, a novel multi-task forecasting model is proposed to take advantages of correlations on different timescales, and to simultaneously report future irradiance values on multiple timescales.

In the proposed MTS model, the main idea is to predict the solar irradiance on different timescale in parallel. The timescale of each task is depending on the collected solar data format. For example, if the solar irradiance is collected in the hourly interval, the proposed model can fulfill multiple tasks of hour-ahead, day-ahead, and week-ahead prediction in parallel, but it cannot conduct the minute-ahead prediction due to information missing. In this article, based on the hourly irradiance data, two tasks at hourly and daily scale are adopted to illustrate the principle of the proposed method.

Following statistical methods in (1), the ARX-based hourly and daily MTS model are formulated as follows:

$$\hat{H}_i = a_{11}H_{i-1} + \cdots + a_{1n_h}H_{i-n_h}$$
$$+ b_{11}D_{i-1} + \cdots + b_{1n_d}D_{i-n_d}$$
$$+ c_{11}W_{i-1} + \cdots + c_{1n_w}W_{i-n_w} + e_h \quad (4)$$

$$\hat{D}_i = a_{21}H_{i-1} + \cdots + a_{2n_h}H_{i-n_h}$$
$$+ b_{21}D_{i-1} + \cdots + b_{2n_d}D_{i-n_d}$$
$$+ c_{21}W_{i-1} + \cdots + c_{2n_w}W_{i-n_w} + e_d \quad (5)$$
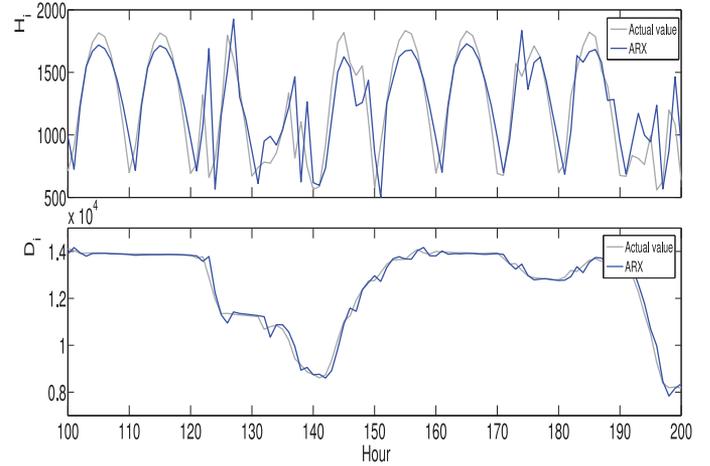
Fig. 1. Hourly (first row) and daily (second row) prediction results by ARX versus actual value.

where n_h, n_d, and n_w denote the length of relative solar irradiance on different timescales, H_i, D_i, and W_i are hourly, daily and weekly solar irradiance. As shown in (4) and (5), the value of solar irradiance at a certain timescale is predicted by those solar irradiance values from different timescales. In this article, the daily and weekly solar irradiance value are denoted as

$$D_{i-k} = \sum_{j=i-24k}^{i-24\,k+23} H_j \quad (6)$$

$$W_{i-k} = \sum_{j=i-168k}^{i-168\,k+167} H_j. \quad (7)$$

Then, the linear multi-input and multi-output (MIMO) ARX model can be generalized as

$$\Lambda = OM + E \quad (8)$$

where $\Lambda = (\hat{H}_i, \hat{D}_i)^T$, $E = (e_h, e_d)^T$ and

$$O = \begin{bmatrix} \underbrace{a_{11} \quad \ldots \quad a_{1n_h}}_{\text{hourly}} & \underbrace{b_{11} \ldots b_{1n_d}}_{\text{daily}} & \underbrace{c_{11} \ldots c_{1n_w}}_{\text{weekly}} \\ a_{21} \ldots a_{2n_h} & b_{21} \ldots b_{2n_d} & c_{21} \ldots c_{2n_w} \end{bmatrix} \quad (9)$$

$$M = [H_{i-1}, \ldots, H_{i-n_h}, D_{i-1}, \ldots, D_{i-n_d}, W_{i-1}, \ldots, W_{i-n_w}]^T. \quad (10)$$

Finally, based the historical data, coefficients O can be calculated to fit the relationship between the predicted and historical values. A test example is used to verify the feasibility of MTS ARX, and the simulation results are shown in Fig. 1. It can be noted that ARX can handle the daily forecast task that has slow dynamics, while ARX has limited ability to handle hourly task that has fast dynamics.

To enhance nonlinear fitting capability, the ANN-based MTS model is established. According to (2), the hourly and daily solar irradiance can be separately forecasted in traditional ANN
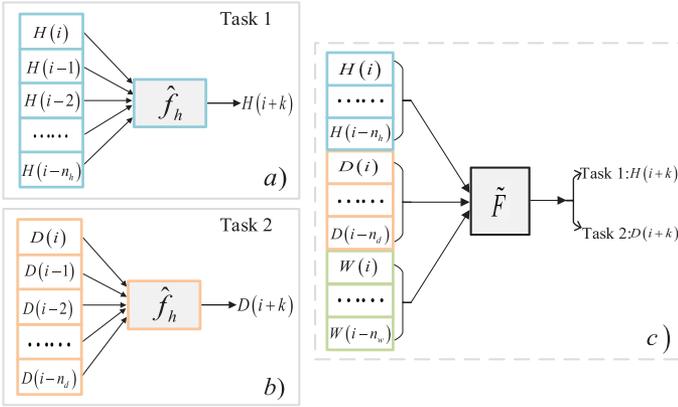
Fig. 2. Single-task (left column) and multi-task (right column) ANN models for solar irradiance forecast.

models as follows:

$$\begin{cases} \hat{H}_i = \tilde{f}_h(H_{i-1}, H_{i-2}, \ldots, H_{i-n_h}) \\ \hat{D}_i = \tilde{f}_d(D_{i-1}, D_{i-2}, \ldots, D_{i-n_d}) \end{cases}. \quad (11)$$

With the perspective of multi-task learning, the ANN-based MTS model is proposed as:

$$\Lambda = \tilde{F}(H_{i-1}, \ldots, H_{i-n_h}, D_{i-1}, \ldots, D_{i-n_d}, W_{i-1}, \ldots, W_{i-n_w}) \quad (12)$$

where $\Lambda = (\hat{H}_i, \hat{D}_i)^T$. Fig. 2 illustrates the difference between single-timescale single-task forecasting model (left column) and multi-timescale multi-task prediction model (right column). For the single-timescale models in Fig. 2(a) and (b), each task only makes use of historical data on a specific timescale, although historical data on other timescale is correlated with the task. For the MTS model shown in Fig. 2(c), the completeness of one task may be helpful to fulfill other tasks, since each task in the multi-task model has shared information on multi-timescale, and taken advantages of correlations with other tasks.

IV. ESN-BASED MTS FORECAST MODEL

In this section, an MTS-ESN algorithm is put forward for the proposed multi-timescale forecast model. The MTS-ESN has integrated the multi-task learning perspective, so it is able to fit multiple complex irradiance curves. Our experimental results could validate the effectiveness of the proposed multi-task learning perspective and MTS-ESN algorithm.

A. Echo State Network

ESN is a kind of recurrent ANN that is composed by an input layer, a hidden recurrent layer (referred to as a dynamical reservoir), and an output layer. In ESN, the input signal, reservoir states, and the output signal are denoted as $\mathbf{u}(i)$, $\mathbf{x}(i)$, $\mathbf{y}(i)$, respectively.

The dynamical reservoir plays a core role in the ESN. It contains a large number of sparse connected neurons, whose connection weights are randomly initialized and remain unchanged during the process of training and testing. On the one hand, the reservoir encodes the temporal information of input signals from low dimensional input space to high dimensional state space; On the other hand, the reservoir provides a complex nonlinear transformation of input signals. Then, supervised learning mechanism is adopted to train readout connection weights in the output layer. The training of ESN only consists of updating the connection weights from reservoir to output layer. Therefore, the training process is simple and very fast compared with classical recurrent ANNs.

Assume that ESN has K, N, and L neurons in the input, hidden, and output layer, respectively. $\mathbf{W}^{\mathrm{in}} \in \mathbb{R}_{N \times K}$, $\mathbf{W}^{\mathrm{res}} \in \mathbb{R}_{N \times N}$ and $\mathbf{W}^{\mathrm{out}} \in \mathbb{R}_{L \times N}$ represent the input-hidden, hidden-hidden and hidden-output connection weight matrices, respectively. The update equations of reservoir states and network outputs are expressed as follows:

$$\mathbf{x}(i+1) = \tanh\left(\mathbf{W}^{\mathrm{in}}\mathbf{u}(i+1) + \mathbf{W}^{\mathrm{res}}\mathbf{x}(i)\right) \quad (13)$$

$$\mathbf{y}(i+1) = \mathbf{W}^{\mathrm{out}}\mathbf{x}(i+1) \quad (14)$$

where tanh denotes the hyperbolic tangent nonlinear function.

In ESN, the main effort is to train the output connection weight $\mathbf{W}^{\mathrm{out}}$. During the learning stage, ESN is trained according to the teacher-forced signal. Simultaneously, the reservoir states updated by (13) are collected in a matrix $\mathbf{X}$. Let l_{tr} represent the length of training data sets, then the reservoir states matrix is denoted as

$$\mathbf{X} = \begin{bmatrix} x_1(1) & x_2(1) & \ldots & x_N(1) \\ x_1(2) & x_2(2) & \ldots & x_N(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(l_{\mathrm{tr}}) & x_2(l_{\mathrm{tr}}) & \ldots & x_N(l_{\mathrm{tr}}) \end{bmatrix}_{l_{\mathrm{tr}} \times N} \quad (15)$$

and the corresponding teacher signal vector matrix is denoted as

$$\mathbf{D} = \begin{bmatrix} d_1(1) & d_2(1) & \ldots & d_L(1) \\ d_1(2) & d_2(2) & \ldots & d_L(2) \\ \vdots & \vdots & \vdots & \vdots \\ d_1(l_{\mathrm{tr}}) & d_2(l_{\mathrm{tr}}) & \ldots & d_L(l_{\mathrm{tr}}) \end{bmatrix}_{l_{\mathrm{tr}} \times L} . \quad (16)$$

The ridge regression training mechanism written in (17) is adopted in this article, in order to overcome the over-fitting phenomenon.

$$(\mathbf{W}^{\mathrm{out}})^T = (\mathbf{X}^T\mathbf{X} + \rho\mathbf{I})^{-1}\mathbf{X}^T\mathbf{D} \quad (17)$$

$\mathbf{I}$ is the N-order identity matrix, and ρ denotes the regularization parameter which should be determined through a large number of experiments for given learning tasks. After the output weights matrix $\mathbf{W}^{\mathrm{out}}$ is obtained, the ESN is ready for testing. The training procedure of ESN is shown in Algorithm 1.

B. Novel MTS-ESN

In the MTS model, a novel ESN structure with the ability of multi-task learning is established, as shown in Fig. 3. For simplicity, in this article we study benchmarks with three input signals and two output signals. That means that historical signals on three different timescales are adopted to predict two different tasks. The input signals are historical hourly, daily and weekly irradiance values. The output signals are the hourly and daily solar irradiance values in the future intervals. In this proposed

---

Algorithm 1: Training Procedures of ESN.

Input: u
Output: $\mathbf{W}^{\text{out}}$

1:    Data preprocessing: divide the collected data into training and testing sets;
2:    Network initialization: initialize an untrained ESN $\{\mathbf{W}^{\text{in}}, \mathbf{W}^{\text{res}}\}$, and set the number of input units K, internal units N, output units L and the spectral radius;
3:    Compute and collect internal states and corresponding target outputs based on (13), (15) and (16) based on the training samples;
4:    Calculate the output weights $\mathbf{W}^{\text{out}}$ according to (17).
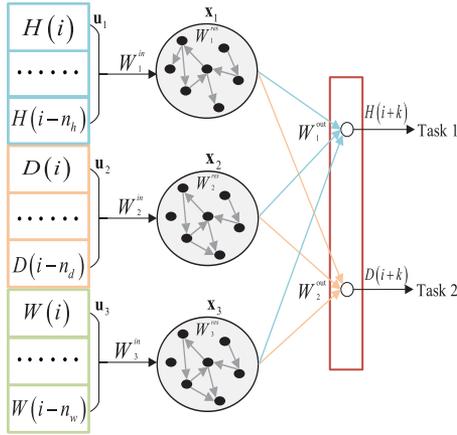
Fig. 3.   Structure of MTS-ESN.

MTS-ESN model, each task interacts with information sharing through multiple reservoirs.

As shown in Fig. 3, the number of reservoirs is the same with the number of input signals on different timescales, so three reservoirs exist in the MTS-ESN model. These three reservoirs are used to map the historical hourly, daily and weekly records to their feature space, separately. Then, these three reservoirs are aggregated at the output layer to predict the future hourly and daily solar irradiance.

Remark 1: When the number of reservoir in MTS-ESN is equal to 1, and a single-task is conducted in the output layer, the resulted structure is the same with the classical ESN. The single-task ESN can be taken as a benchmark, to show the contribution of the correlations between each task to the prediction accuracy of MTS-ESN in later simulation.

Denote input vectors as $\mathbf{u}_1(i) = [H(i), H(i-1), \ldots, H(i - K_1)]^T$, $\mathbf{u}_2(i) = [D(i), D(i-1), \ldots, D(i - K_2)]^T$, and $\mathbf{u}_3(i) = [W(i), W(i-1), \ldots, W(i - K_3)]^T$, where K_1, K_2 and K_3 are the length of each input signal, i.e., the length of historical hourly irradiance values n_h, daily irradiance values n_d, and weekly irradiance values n_w in Fig. 3, respectively. The internal states of three reservoirs are $\mathbf{x}_1(i) = [x_1^1(i), x_1^2(i), \ldots, x_1^{N_1}(i)]^T$, $\mathbf{x}_2(i) = [x_2^1(i), x_2^2(i), \ldots, x_2^{N_2}(i)]^T$, and $\mathbf{x}_3(i) = [x_3^1(i), x_3^2(i), \ldots, x_3^{N_3}(i)]^T$, where N_1, N_2 and N_3 are the

number of internal units of three reservoirs. The output vector is $\mathbf{y} = [\mathbf{y}_1; \mathbf{y}_2]$, where $\mathbf{y}_1 = [H(i+1), \ldots, H(i + L_1)]$ and $\mathbf{y}_2 = [D(i+1), \ldots, D(i + L_2)]$. L_1 and L_2 are the length of output single for each task.

$\mathbf{W}_l^{\text{in}} \in \mathbb{R}^{N_l \times K_l}$ and $\mathbf{W}_l^{\text{res}} \in \mathbb{R}^{N_l \times N_l}$ represent the input-hidden, hidden-hidden connection weight matrices for reservoir l ($l = 1, 2, 3$). $\mathbf{W}_1^{\text{out}} \in \mathbb{R}^{L_1 \times (N_1 + N_2 + N_3)}$ and $\mathbf{W}_2^{\text{out}} \in \mathbb{R}^{L_2 \times (N_1 + N_2 + N_3)}$ represent the hidden-output connection weight matrices for each task. In particular, we study a one-step multi-task forecasting model, i.e., $L_1 = 1$ and $L_2 = 1$.

Based on the input signals and initialized weights, the reservoir states at time $i + 1$ can be computed as:

$$\begin{cases} \mathbf{x}_1(i + 1) = \tanh\left(\mathbf{W}_1^{\text{in}}\mathbf{u}_1(i + 1) + \mathbf{W}_1^{\text{res}}\mathbf{x}_1(i)\right) \\ \mathbf{x}_2(i + 1) = \tanh\left(\mathbf{W}_2^{\text{in}}\mathbf{u}_2(i + 1) + \mathbf{W}_2^{\text{res}}\mathbf{x}_2(i)\right) \\ \mathbf{x}_3(i + 1) = \tanh\left(\mathbf{W}_3^{\text{in}}\mathbf{u}_3(i + 1) + \mathbf{W}_3^{\text{res}}\mathbf{x}_3(i)\right) \end{cases} \quad (18)$$

Then, the network outputs at time $i + 1$ are expressed as follows:

$$\begin{cases} \mathbf{y}_1(i + 1) = \mathbf{W}_1^{\text{out}}[\mathbf{x}_1(i + 1); \mathbf{x}_2(i + 1); \mathbf{x}_3(i + 1)] \\ \mathbf{y}_2(i + 1) = \mathbf{W}_2^{\text{out}}[\mathbf{x}_1(i + 1); \mathbf{x}_2(i + 1); \mathbf{x}_3(i + 1)]. \end{cases} \quad (19)$$

Let $\mathbf{y}(i + 1) = [\mathbf{y}_1(i + 1); \mathbf{y}_2(i + 1)] \in \mathbb{R}^{(L_1 + L_2) \times 1}$, $\mathbf{x}(i + 1) = [\mathbf{x}_1(i + 1); \mathbf{x}_2(i + 1); \mathbf{x}_3(i + 1)] \in \mathbb{R}^{(N_1 + N_2 + N_3) \times 1}$, and $\mathbf{W}^{\text{out}} = [\mathbf{W}_1^{\text{out}}; \mathbf{W}_2^{\text{out}}] \in \mathbb{R}^{(L_1 + L_2) \times (N_1 + N_2 + N_3)}$. Then, we can obtain that

$$\begin{cases} y_1(i + 1) = \mathbf{W}_1^{\text{out}}\mathbf{x}(i + 1) \\ y_2(i + 1) = \mathbf{W}_2^{\text{out}}\mathbf{x}(i + 1) \end{cases} \quad (20)$$

and that

$$\mathbf{y}(i + 1) = \mathbf{W}^{\text{out}}\mathbf{x}(i + 1). \quad (21)$$

Similar to traditional ESN, the ridge regression training mechanism can also be adopted to compute the output weights. Equation (17) is rewritten as:

$$(\mathbf{W}^{\text{out}})^T = (\mathbf{Q}^T\mathbf{Q} + \rho\mathbf{I})^{-1}\mathbf{Q}^T\mathbf{T} \quad (22)$$

where $\mathbf{I}$ is the $(N_1 + N_2 + N_3)$-order identity matrix. Let l_{tr} represent the length of training samples, then the reservoir states matrix $\mathbf{Q}$ and corresponding teacher signal vector matrix $\mathbf{T}$ are denoted as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{x}_1^T(1) & \mathbf{x}_2^T(1) & \mathbf{x}_3^T(1) \\ \mathbf{x}_1^T(2) & \mathbf{x}_2^T(2) & \mathbf{x}_3^T(2) \\ \vdots & \vdots & \vdots \\ \mathbf{x}_1^T(l_{\text{tr}}) & \mathbf{x}_2^T(l_{\text{tr}}) & \mathbf{x}_N^T(l_{\text{tr}}) \end{bmatrix}_{l_{\text{tr}} \times (N_1 + N_2 + N_3)} \quad (23)$$

and

$$\mathbf{T} = \begin{bmatrix} d_1(1) & d_2(1) \\ d_1(2) & d_2(2) \\ \vdots & \vdots \\ d_1(l_{\text{tr}}) & d_2(l_{\text{tr}}) \end{bmatrix}_{l_{\text{tr}} \times L} \quad (24)$$

$\mathbf{d}_1$ and $\mathbf{d}_2$ are the teacher signal during training stage, i.e., the hourly and daily solar irradiance to be predicted. The training procedure of MTS-ESN for solar irradiance is shown in Algorithm 2.

Remark 2: In the proposed MTS-ESN model, correlations among solar irradiance from different timescales are utilized for

---

Algorithm 2: Training Procedures of MTS-ESN.

Input: $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$

Output: $\mathbf{W}^{\text{out}}$

1: Data preprocessing: divide the original collected hourly solar irradiance data into daily and weekly data sets according to (6)-(7), and then split these data sets into training and testing sets according to Table I;

2: **for** $l \leftarrow 1$ to 3

3: Initialization: $\{\mathbf{W}_l^{\text{in}}, \mathbf{W}_l^{\text{res}}, K_l, N_l, \rho\}$;

4: **end**

5: **for** $i \leftarrow 1$ to l_{tr}

6: **for** $l \leftarrow 1$ to 3

7:   $\mathbf{x}_l(i+1) \leftarrow \tanh(\mathbf{W}_l^{\text{in}}\mathbf{u}_l(i+1) + \mathbf{W}_l^{\text{res}}\mathbf{x}_l(i))$;

8: **end**

9: $\mathbf{d}(i) \leftarrow [H(i), D(i)]$;

10: $\mathbf{x}(i) \leftarrow [\mathbf{x}_1(i); \mathbf{x}_2(i); \mathbf{x}_3(i)]$;

11: **end**

12: $\mathbf{Q} \leftarrow (\mathbf{x}(1); \mathbf{x}(2); \ldots; \mathbf{x}(l_{\text{tr}}))$;

13: $\mathbf{T} \leftarrow (\mathbf{d}(1); \mathbf{d}(2); \ldots; \mathbf{d}(l_{\text{tr}}))$;

14: $(\mathbf{W}^{\text{out}})^T \leftarrow (\mathbf{Q}^T\mathbf{Q} + \rho\mathbf{E})^{-1}\mathbf{Q}^T\mathbf{T}$;

---

multi-task modeling. Each task in MTS-ESN interacts with information sharing through multiple reservoirs. Due to the simple training mechanism and information sharing among each task, MTS-ESN can fulfill multiple tasks in parallel more efficiently and and accurately than multiple single-task ESN models. It should be noted that the correlation among tasks is crucial for multi-task learning, as unrelated tasks will transmit noisy information. In this case, single-task models are better than multi-task models.

V. EXPERIMENTAL DESIGN AND RESULTS

A. Performance Evaluation Matrices

In this article, four different metrics in [1] are adopted to evaluate the accuracy of MTS-ESN, i.e., root mean square error (RMSE), mean absolute error (MAE), normalized root mean square error (nRMSE), and correlation coefficient (R). These metrics are expressed in the following equations:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{\text{forecasted},i} - y_{\text{measured},i}\right)^2} \quad (25)$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_{\text{forecasted},i} - y_{\text{measured},i}\right| \quad (26)$$

$$\text{nRMSE} = \frac{1}{\bar{y}}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{\text{forecasted},i} - y_{\text{measured},i}\right)^2} \times 100\% \quad (27)$$

$$R = \frac{\text{cov}\left(y_{\text{forecasted}}, y_{\text{measured}}\right)}{\sigma_{y_{\text{forecasted}}}\sigma_{y_{\text{measured}}}} \quad (28)$$

where $y_{\text{forecasted},i}$ is the forecasted solar irradiance, $y_{\text{measured},i}$ is the measured solar irradiance, $\bar{y}$ is the mean of actual irradiance
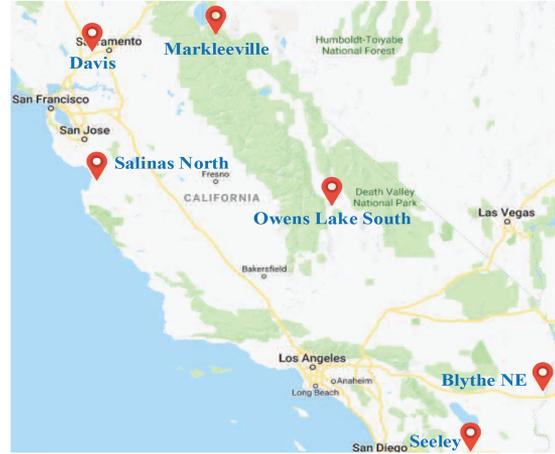
Fig. 4. Map information of six different stations.

TABLE I
TRAINING AND TESTING PERIOD FOR PREDICTION MODELS

Case Number	Training period	Testing period
I	Jan.1∼Feb.28	Mar.1∼Mar.31
II	Apr.1∼May.31	Jun.1∼Jun.30
III	Jul.1∼Aug.31	Sep.1∼Sep.30
IV	Oct.1∼Nov.30	Dec.1∼Dec.31

values at the test period, and n is the length of testing step. $\text{cov}(y_{\text{forecasted}}, y_{\text{measured}})$ is the covariance between forecasted and measured values, $\sigma_{y_{\text{forecasted}}}$ and $\sigma_{y_{\text{measured}}}$ are the standard deviations of $y_{\text{forecasted}}$ and y_{measured}.

B. Data Description

In this article, the hourly solar irradiance records of 6 different stations at California are obtained from California Irrigation Management Information System (CIMIS). The six stations are Davis, Markleeville, Owens Lake South, Salinas North, Seeley and Blythe NE, as shown in Fig. 4. Since the solar irradiance before sunrise and after sunset can be negligible, the data from 8 A.M. to 5 P.M. in one day are therefore chosen in the current study. That is to say, the length of a supposed day equals to 10. For any station, the training and testing period are shown in Table I.

C. Performance of MTS-ESN

In this section, the prediction performance of MTS-ESN is evaluated in the hourly and daily forecasting. The network parameters are selected by trial and error, and they are set as follows: each reservoir size is set as 200, the total input number is 6 ($K_1 = 2$, $K_2 = 2$ and $K_3 = 2$), and the output number is set as 2; Weight matrix $\mathbf{W}^{\text{in}}$ is sampled from a uniform distribution over $[-1,1]$, and the spectral radius of $\mathbf{W}^{\text{res}}$ is set to be 0.85; The regularization parameter ρ of ridge regression training algorithm is set as 10^{-1}.
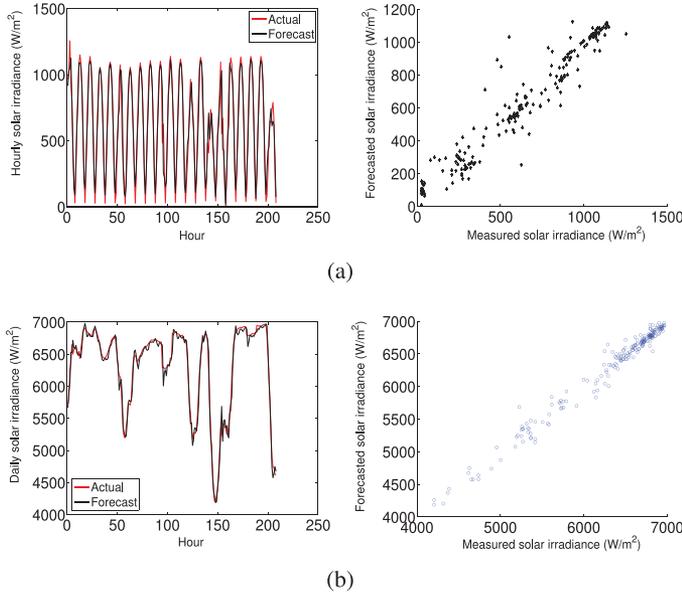
Fig. 5.   (a) Hourly and (b) daily prediction results by MTS-ESN versus actual values for CASE IV at Blythe NE station.

TABLE II
HOURLY STATISTICAL TEST COMPARISON BETWEEN SINGLE-TASK ESN MODEL AND MULTITASK MTS-ESN MODEL FOR CASE I AT DAVIS STATION AND CASE IV AT BLYTHE NE STATION

Case	Model	RMSE (W/m^2)	nRMSE (%)	MAE (W/m^2)	R
I	ESN	232.44	23.31	166.86	0.892
	MTS-ESN	220.98	22.16	169.68	0.903
IV	ESN	95.42	15.12	64.02	0.964
	MTS-ESN	93.19	14.76	64.71	0.966

TABLE III
DAILY STATISTICAL TEST COMPARISON BETWEEN SINGLE-TASK ESN MODEL AND MULTITASK MTS-ESN MODEL FOR CASE I AT OWENS LAKE SOUTH STATION AND CASE IV AT BLYTHE NE STATION

Case	Model	RMSE (W/m^2)	nRMSE (%)	MAE (W/m^2)	R
I	ESN	350.67	3.57	288.17	0.993
	MTS-ESN	282.50	2.87	228.26	0.995
IV	ESN	109.91	1.74	74.43	0.988
	MTS-ESN	97.62	1.55	68.55	0.990

To show the contribution of the correlation between each task to the prediction accuracy, two single-task ESN model and a multi-task MTS-ESN model are firstly conducted, based on the historical signals on three different timescales. Take CASE I at Davis station and CASE IV at Blythe NE station as two examples. The hourly and daily statistical test results in terms of RMSE, nRMSE, MAE, and *R* are compared in Tables II and III, respectively. Obviously, the multi-task MTS-ESN model performs better than single-task ESN model, at both hourly and daily levels. The results illustrate that the correlation between each forecasting task contributes to the prediction accuracy improvement of MTS-ESN.

Fig. 5 shows the hourly and daily forecasted value of MTS-ESN versus the actual value for CASE IV at Blythe NE station, respectively. 200 time points in the testing stages are selected.
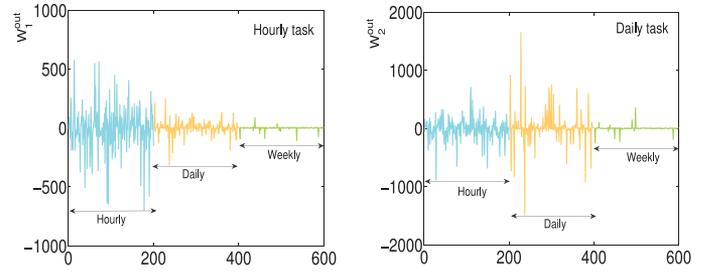
Fig. 6.   Output weights distribution comparison for hourly and daily prediction.

TABLE IV
HOURLY STATISTICAL TEST BETWEEN MEASURED AND FORECASTED VALUES FOR FOUR CASES AT SIX STATIONS

Station	Case	RMSE (W/m^2)	nRMSE (%)	MAE (W/m^2)	R
Davis	I	220.98	22.16	169.68	0.903
	II	196.95	12.71	82.91	0.878
	III	168.74	15.62	123.68	0.924
	IV	93.67	19.89	71.98	0.954
Markleeville	I	288.16	27.89	208.53	0.804
	II	206.81	13.70	124.98	0.896
	III	252.81	24.41	165.11	0.883
	IV	138.79	28.20	102.14	0.914
Owens Lake South	I	271.65	21.09	191.40	0.860
	II	139.79	8.25	82.70	0.948
	III	179.93	13.95	115.79	0.932
	IV	105.48	14.91	75.28	0.970
Salinas North	I	221.32	22.96	166.09	0.896
	II	178.75	13.17	124.08	0.936
	III	127.65	13.78	97.12	0.944
	IV	79.57	14.44	59.18	0.971
Seeley	I	235.26	18.83	183.06	0.919
	II	69.63	4.38	49.79	0.986
	III	107.21	8.43	73.75	0.974
	IV	195.06	32.48	139.29	0.831
Blythe NE	I	166.21	15.08	133.74	0.953
	II	69.74	4.60	51.73	0.985
	III	87.21	7.49	59.05	0.980
	IV	93.19	14.76	64.71	0.966

From Fig. 5, it can be observed that the MTS-ESN model can not only fit hourly solar irradiance, but also fit the daily solar irradiance data.

Fig. 6 shows the corresponding trained output weights distribution at three timescales. It can be seen that for hourly and daily prediction tasks, the output weights are concentrated in the corresponding timescale.

For each case, the prediction accuracy of MTS-ESN is evaluated and computed in terms of RMSE, MAE, nRMSE and *R* for each station, as reported in Tables IV and V. According to the statistical results obtained, it can be observed that nRMSE varies in the range of 0.36% − 3.33% for daily prediction results, and 4.38% − 32.48% for hourly simulation results. The statistical results also illustrate that the prediction task at daily timescale is easier than hourly timescale, as the daily dynamics is slower than the hourly dynamics. In sum, the developed MTS-ESN model is suitable for the simultaneous prediction of solar irradiance at hourly and daily timescale.
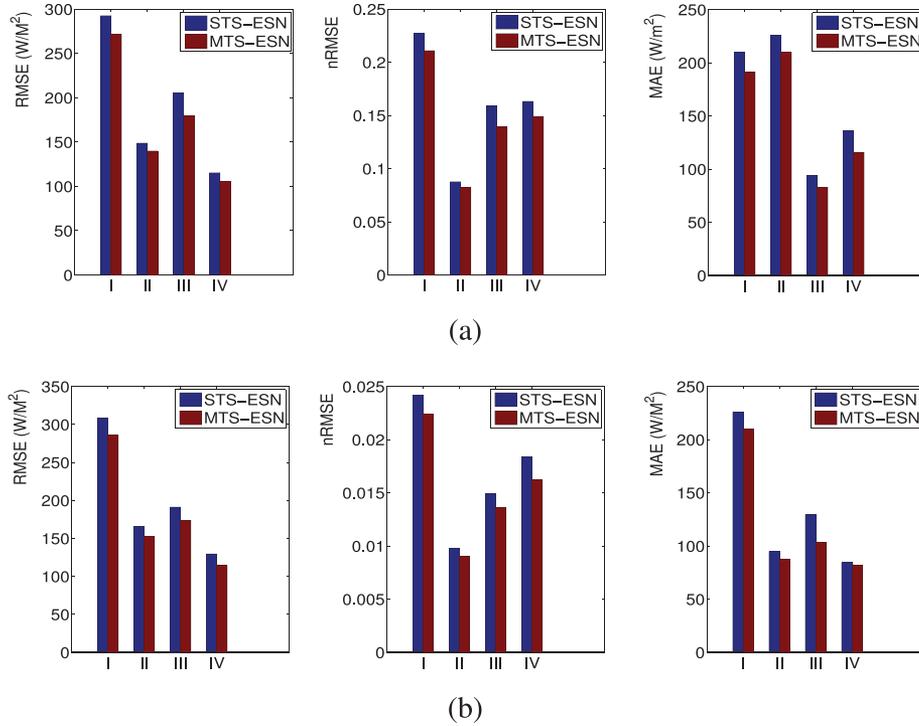
Fig. 7. (a) Hourly and (b) daily comparison between the measured and predicted solar irradiance by MTS-ESN and STS-ESN models for four cases at Owens Lake South station.

TABLE V
DAILY STATISTICAL TEST BETWEEN MEASURED AND FORECASTED
VALUES FOR FOUR CASES AT SIX STATIONS

Station	Case	RMSE (W/m^2)	nRMSE (%)	MAE (W/m^2)	R
Davis	I	282.50	2.87	228.26	0.995
	II	194.52	1.26	89.99	0.987
	III	176.90	1.64	130.09	0.975
	IV	97.38	2.06	75.89	0.995
Markleeville	I	341.93	3.33	262.85	0.984
	II	224.22	1.49	132.19	0.992
	III	268.08	2.58	186.24	0.989
	IV	148.88	3.03	111.39	0.987
Owens Lake South	I	285.81	2.24	209.88	0.981
	II	152.14	0.90	87.76	0.988
	III	173.93	1.36	103.21	0.989
	IV	114.41	1.62	81.88	0.988
Salinas North	I	276.91	2.91	210.55	0.995
	II	196.78	1.45	151.21	0.995
	III	133.17	1.44	99.72	0.994
	IV	75.41	1.37	56.61	0.997
Seeley	I	349.94	2.82	302.14	0.989
	II	71.90	0.45	48.81	0.993
	III	103.08	0.81	73.53	0.966
	IV	193.31	3.22	140.34	0.977
Blythe NE	I	270.30	2.46	239.02	0.986
	II	53.69	0.36	38.59	0.993
	III	91.03	0.78	70.50	0.946
	IV	97.62	1.55	68.55	0.990

TABLE VI
STATISTICAL TEST COMPARISON BETWEEN MEASURED AND FORECASTED
VALUES BY MTS-ESN AND STS-ESN MODELS FOR FOUR
CASES AT OWENS LAKE SOUTH STATION

Case	Model	RMSE (W/m^2)	nRMSE (%)	MAE (W/m^2)	R	Time(s)
I	STS-ESN(h)	292.69	22.72	209.52	0.843	2.132
	MTS-ESN(h)	271.65	21.09	191.40	0.860	1.485
	STS-ESN(d)	308.35	2.42	225.92	0.980	2.122
	MTS-ESN(d)	285.81	2.24	209.88	0.981	1.485
II	STS-ESN(h)	148.71	8.77	93.91	0.943	2.058
	MTS-ESN(h)	139.79	8.25	82.70	0.948	1.480
	STS-ESN(d)	165.10	0.98	94.85	0.986	2.094
	MTS-ESN(d)	152.14	0.90	87.76	0.988	1.480
III	STS-ESN(h)	205.49	15.93	136.26	0.912	2.072
	MTS-ESN(h)	179.93	13.95	115.79	0.932	1.379
	STS-ESN(d)	191.17	1.49	129.52	0.988	2.124
	MTS-ESN(d)	173.93	1.36	103.21	0.989	1.379
IV	STS-ESN(h)	115.47	16.32	81.66	0.965	2.047
	MTS-ESN(h)	105.48	14.91	75.28	0.970	1.438
	STS-ESN(d)	129.48	1.84	84.31	0.985	2.041
	MTS-ESN(d)	114.41	1.62	81.88	0.988	1.438

(STS-ESN) shown in Fig. 2(a) and (b). For STS-ESN, the hourly and daily forecast models are constructed, respectively. In order to ensure the fairness of experiments, the parameters of STS-ESN are set to be same with MTS-ESN. Take Owens Lake South station as an example. The statistical results in terms of RMSE, MAE, nRMSE, *R* and computation time are shown in Fig. 7 and Table VI. In Table VI, STS-ESN(h) and MTS-ESN(h) denote the hourly forecast results by STS-ESN and MTS-ESN, respectively. STS-ESN(d) and MTS-ESN(d) denote the daily forecast results by STS-ESN and MTS-ESN, respectively. It is

D. Comparisons of MTS-ESN With STS-ESN

In addition, the performance of MTS-ESN is further evaluated by comparing with that of single-timescale ESN model

TABLE VII
HOURLY STATISTICAL TEST COMPARISON BETWEEN DIFFERENT
MODELS FOR CASE I AT DAVIS STATION AND CASE III
AT OWENS LAKE SOUTH STATION

Case	Model	RMSE (W/m^2)	nRMSE (%)	MAE (W/m^2)	R	Time(s)
	PSS	310.87	31.18	254.87	0.817	1.536
	ARX	239.09	23.98	186.79	0.886	4.617
I	MTS-BP	383.37	38.45	291.29	0.769	3.693
	MTS-Elman	238.78	23.95	179.87	0.888	5.453
	MTS-ESN	**220.98**	**22.16**	**169.68**	**0.903**	**2.911**
	PSS	336.69	26.10	292.08	0.744	1.479
	ARX	227.12	17.61	177.37	0.878	4.556
III	MTS-BP	220.28	17.08	158.96	0.893	3.920
	MTS-Elman	233.01	18.06	173.80	0.879	5.453
	MTS-ESN	**179.93**	**13.95**	**115.79**	**0.932**	**2.757**

TABLE VIII
DAILY STATISTICAL TEST COMPARISON BETWEEN DIFFERENT MODELS FOR
CASE I AT DAVIS STATION AND CASE III AT OWENS LAKE SOUTH STATION

Case	Model	RMSE (W/m^2)	nRMSE (%)	MAE (W/m^2)	R	Time(s)
	MTS-BP	394.86	4.02	310.35	0.989	3.693
I	MTS-Elman	338.75	3.45	257.65	0.992	5.453
	MTS-ESN	**282.50**	**2.87**	**228.26**	**0.995**	**2.911**
	MTS-BP	299.99	2.34	212.04	0.977	3.920
III	MTS-Elman	249.72	1.95	176.81	0.986	5.453
	MTS-ESN	**173.93**	**1.36**	**103.21**	**0.989**	**2.757**

worth noting that the prediction accuracy of MTS-ESN is higher than STS-ESN, at both hourly and daily levels. Furthermore, the computation time of STS-ESN and MTS-ESN are reported. Note that the computation time of MTS-ESN(h) and MTS-ESN(d) is the average time, as these two values are obtained simultaneously in the overall computation of MTS-ESN. Experimental results indicate that the MTS-ESN has much less computation time than STS-ESN for each task. With respect to prediction accuracy and computational efficiency, MTS-ESN performs better than STS-ESN to deal with multi-task prediction of solar irradiance. The reason is that the information in one task could help to fulfill other tasks.

E. Comparisons of MTS-ESN and Other Typical Models

In order to further validate the effectiveness of the proposed MTS model, the performance of MTS-ESN, multi-timescale Elman (MTS-Elman) and multi-timescale BP (MTS-BP) networks are compared in this section. In addition, the classical PSS and ARX are taken as two comparative benchmarks for hourly solar irradiance prediction task.

Take CASE I at Davis station and CASE III at Owens Lake South as two examples. Comparative results in terms of RMSE, nRMSE, MAE, computation time and R are listed in Tables VII and VIII.

It is obvious that the multi-task learning perspective still works in MTS-BP and MTS-Elman. MTS-ESN performs better than MTS-Elman and MTS-BP at both hourly and daily levels, with higher prediction accuracy and less computation time. For the hourly task, MTS-ESN also outperforms than PSS and ARX benchmarks. Those time slots in the testing stages are sampled to evaluate the results between actual and predicted values, as
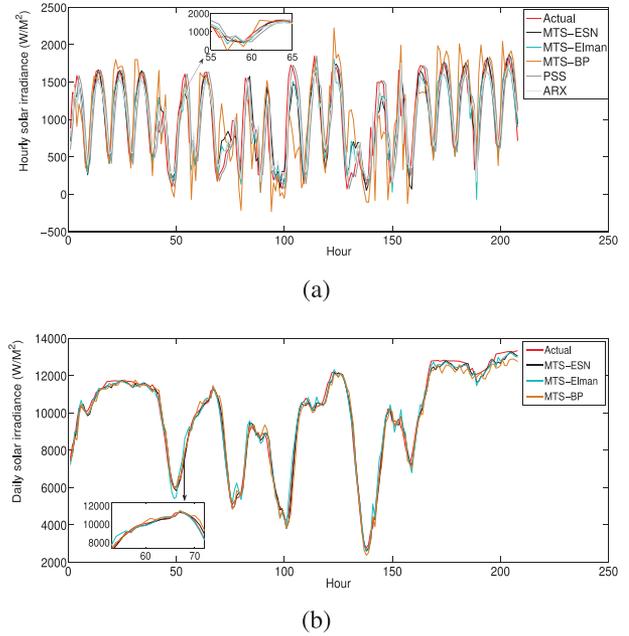
Fig. 8. (a) Hourly and (b) daily comparison between the measured and predicted solar irradiance by different models for Case I at Davis station.

shown in Fig. 8. From Fig. 8(a), it can be seen that there exist large fluctuations in the forecasted hourly values by MTS-BP, ARX and MTS-Elman models at the sampling slots, compared with the MTS-ESN model. From the daily comparison between the measured and predicted solar irradiance by three ANN models shown in Fig. 8(b), it can be observed that the accuracy of daily prediction increase for three models, compared with that of hourly prediction. The reason is that the daily solar irradiance has slower dynamics than hourly solar irradiance. However, the MTS-ESN model could also perform better than MTS-BP and MTS-Elman models. As a result, it can be concluded that the proposed multi-task learning perspective could be applied to different ANNs, and that MTS-ESN algorithm has better performance to deal with multi-timescale solar irradiance prediction task than the MTS-BP and MTS-Elman algorithms.

VI. CONCLUSION

In this article, a new multi-task learning perspective is proposed to fulfill the MTS forecast of solar irradiance. A novel MTS-ESN algorithm is proposed to verify the usefulness of task correlation in the multi-timescale irradiance prediction. In the MTS-ESN, the information on different tasks are shared by integration of multiple reservoirs. Simulation results demonstrate that our proposed MTS-ESN algorithm could perform effectively at both hourly and daily prediction in parallel, which also indicate the effectiveness of multi-task learning in this application.

In the comparative study, the proposed MTS-ESN could outperform the existing STS-ESN with respect to prediction accuracy and computational efficiency. The multi-task learning perspective could be applied to other ANNs, e.g., BP, Elman neural networks. As the MTS-ESN has better performance than

MTS-BP and MTS-Elman for multi-task prediction, we only focus on ESN and MTS-ESN in this article. The outputs of this article simultaneously complete multiple prediction tasks, which could help make hierarchical decisions for energy systems. In a hybrid solar energy system, the outputs can also help customers re-adjust the load demand in advance. In addition, the proposed method can also be applied to existing single-task ANN models and other learning tasks for further performance improvement. In future work, more ANNs could be evaluated in the multi-timescale forecast of solar irradiance.

REFERENCES

[1] C. Voyant *et al.*, "Machine learning methods for solar radiation forecasting: A review," *Renewable Energy*, vol. 105, pp. 569–582, 2017.

[2] M. Diagne *et al.*, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renewable Sustain. Energy Rev.*, vol. 27, no. 6, pp. 65–76, 2013.

[3] J. Boland, "Time series and statistical modelling of solar radiation," in *Recent advances in solar radiation modelling*. Berlin, Germany: Springer-Verlag, 2008, pp. 283–312.

[4] D. V. D. Meer, G. R. C. Mouli, G. M.-E. Mouli, L. R. Elizondo, and P. Bauer, "Energy management system with PV power forecast to optimally charge EVs at the workplace," *IEEE Trans. Ind. Informat.*, vol. 14, no. 1, pp. 311–320, Jan. 2018.

[5] A. Sfetsos and A. H. Coonick, "Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques," *Sol. Energy*, vol. 68, no. 2, pp. 169–178, 2000.

[6] T. Wada *et al.*, "Estimation of hourly global solar irradiation on tilted planes from horizontal one using artificial neural networks," *Energy*, vol. 39, no. 1, pp. 166–179, 2012.

[7] S. Lou *et al.*, "Prediction of diffuse solar irradiance using machine learning and multivariable regression," *Appl. Energy*, vol. 181, pp. 367–374, 2016.

[8] A. Khosravi *et al.*, "Prediction of hourly solar radiation in Abu Musa Island using machine learning algorithms," *J. Cleaner Prod.*, vol. 176, pp. 63–75, 2018.

[9] R. J. Bessa, A. Trindade, and V. Miranda, "Spatial-temporal solar power forecasting for smart grids," *IEEE Trans. Ind. Informat.*, vol. 11, no. 1, pp. 232–241, Feb. 2015.

[10] A. Tascikaraoglu *et al.*, "Compressive Spatio-temporal forecasting of meteorological quantities and photovoltaic power," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1295–1305, Jul. 2016.

[11] S. Ferlito, G. Adinolfi, and G. Graditi, "Comparative analysis of data-driven methods online and offline trained to the forecasting of grid-connected photovoltaic plant production," *Appl. Energy*, vol. 205, pp. 116–129, 2017.

[12] M. Pierro *et al.*, "Data-driven upscaling methods for regional photovoltaic power estimation and forecast using satellite and numerical weather prediction data," *Sol. Energy*, vol. 158, pp. 1026–1038, 2017.

[13] G. Capizzi, C. Napoli, and F. Bonanno, "Innovative second-generation wavelets construction with recurrent neural networks for solar radiation forecasting," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 11, pp. 1805–1815, Nov. 2012.

[14] S. X. Chen, H. B. Gooi, and M. Q. Wang, "Solar radiation forecast based on fuzzy logic and neural networks," *Renewable Energy*, vol. 60, no. 4, pp. 195–201, 2013.

[15] D. Sáez, F. Ávila, D. Olivares, C. Cañizares, and L. Marín, "Fuzzy prediction interval models for forecasting renewable resources and loads in microgrids," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 548–556, Mar. 2015.

[16] Y. Wang, Y. Shen, S. Mao, G. Cao, and R. M. Nelms, "Adaptive learning hybrid model for solar intensity forecasting," *IEEE Trans. Ind. Informat.*, vol. 14, no. 4, pp. 1635–1645, Apr. 2018.

[17] K. Benmouiza and A. Cheknane, "Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models," *Energy Convers. Manage.*, vol. 75, no. 5, pp. 561–569, 2013.

[18] R. Caruana, "Multitask Learning," *Mach. Learn.*, vol. 28, no. 1, pp. 41–75, 1997.

[19] H. Jaeger and H. Haas, "Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, no. 5667, pp. 78–80, 2004.

[20] C. Gallicchio and A. Micheli, "Tree echo state networks," *Neurocomputing*, vol. 101, pp. 319–337, 2013.

[21] A. J. Wootton *et al.*, "Optimizing Echo State Networks for Static Pattern Recognition," *Cogn. Comput.*, vol. 9, no. 3, pp. 391–399, 2017.

[22] G. Shi, D. Liu, and Q. Wei, "Echo state network-based Q-learning method for optimal battery control of offices combined with renewable energy," *IET Control Theory Appl.*, vol. 11, no. 7, pp. 915–922, 2017.

[23] X. J. Wang, Y. C. Jin, and K. R. Hao, "Echo state networks regulated by local intrinsic plasticity rules for regression," *Neurocomputing*, vol. 351, pp. 111–122, 2019.

[24] N. Chouikhi *et al.*, "PSO-based analysis of echo state network parameters for time series forecasting," *Appl. Soft Comput.*, vol. 55, pp. 211–225, 2017.

[25] C. Yang *et al.*, "Online sequential echo state network with sparse RLS algorithm for time series prediction," *Neural Netw.*, vol. 118, pp. 32–42, 2019.

[26] J. Cao and X. Lin, "Application of the diagonal recurrent wavelet neural network to solar irradiation forecast assisted with fuzzy technique," *Eng. Appl. Artif. Intell.*, vol. 21, no. 8, pp. 1255–1263, 2008.

[27] F. O. Hocaoglu and M. Kurban, "A novel 2-D model approach for the prediction of hourly solar radiation," in *Proc. Int. Work-Conf. Artif. Neural Netw.*, vol. 685, pp. 749–756, 2007.

[28] A. Hasni *et al.*, "Estimating global solar radiation using artificial neural network and climate data in the south-western region of Algeria," *Energy*, vol. 18, no. 1, pp. 531–537, 2012.

[29] Q. Li, Z. Wu, and X. H. Xia, "Estimate and characterize PV power at demand-side hybrid system," *Appl. Energy*, vol. 218, pp. 66–77, 2018.

[30] H. Zheng *et al.*, "A multi-task model for simultaneous face identification and facial expression recognition," *Neurocomputing*, vol. 171, pp. 515–523, 2016.

[31] G. Skolidis and G. Sanguinetti, "Bayesian multitask classification with gaussian process priors," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 22, pp. 2011–2021, Dec. 2011.

[32] L. H. Zhang *et al.*, "Deep multi-level networks with multi-task learning for saliency detection," *Neurocomputing*, vol. 312, pp. 229–238, 2018.

[33] R. Chandra, Y. S. Ong, and C. K. Goh, "Co-evolutionary multi-task learning for dynamic time series prediction," *Appl. Soft Comput.*, vol. 70, pp. 576–589, 2018.

[34] J. B. Fiot and F. Dinuzzo, "Electricity demand forecasting by multi-task learning," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 544–551, Mar. 2018.

[35] E. Adeli *et al.*, "Multi-task prediction of infant cognitive scores from longitudinal incomplete neuroimaging data," *NeuroImage*, vol. 185, pp. 783–792, 2019.

[36] T. Shireen *et al.*, "Iterative multi-task learning for time-series modeling of solar panel PV outputs," *Appl. Energy*, vol. 212, pp. 654–662, 2018.

Zhou Wu (Member, IEEE) received the M.S. degree in control engineering from Wuhan University, Hubei, China, in 2009, and the Ph.D. degree in electronic engineering from the City University of Hong Kong Kowloon Tong, Hong Kong, in 2013.

He is currently a Full Professor with Chongqing University (CQU), Chongqing, China. Before joining in CQU, he worked as a Senior Research Fellow with the University of Pretoria, Pretoria, South Africa, from 2012 to 2015. His current research interests include multidisciplinary areas, energy/building information modeling, intelligent optimization/control, and game theory.

Prof. Wu is a member of Association of Computer Machinery and Chinese Association of Automation. He published more than 80 peer-reviewed papers on *Science Advances*, IEEE TRANSACTION ON INDUSTRIAL INFORMATICS, *Applied Energy*, and other top journals. Furthermore, he has engineering experiences on measurement and verification (MV) energy efficiency projects, and holds several relative patents.

Qian Li received the M.S. degree in control science and engineering, from Chongqing University, Chongqing, China, in 2017, where she is currently working toward the Ph.D. degree in control theory and control engineering.

Her current research interests include intelligent control, neural networks, and time series prediction.

Xiaohua Xia (Fellow, IEEE) received the B.Sc. degree in applied mathematics from Wuhan Institute of Hydraulic and Electrical Engineering, Wuhan, China, in 1983 and the Ph.D. degree in automatic control theory and application from Beijing University of Aeronautics and Astronautics, Beijing, China, in 1989.

He is currently a Professor in the Electrical, Electronic, and Computer Engineering Department, University of Pretoria, Pretoria, South Africa, Director of the Center of New Energy Systems, and the Director of the National Hub for the Postgraduate Programme in Energy Efficiency and Demand-Side Management. He was academically affiliated with the University of Stuttgart, Stuttgart, Germany, the Ecole Centrale de Nantes, Nantes, France, and the National University of Singapore, Singapore before joining the University of Pretoria, Pretoria, South Africa in 1998. His current research interests are industrial energy systems and building energy systems.

Prof. Xia is an IEEE fellow and an NRF A-rated scientist. He was elected a fellow of the South African Academy of Engineering in 2005, and a member of the Academy of Science of South Africa in 2011. He has been an associate editor of Automatica, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, IEEE TRANSACTIONS ON AUTOMATIC CONTROL, and specialist editor (control) of the SAIEE Africa Research Journal, and currently sits at the editorial board of Applied Energy.