An interval-valued time series forecasting scheme with probability distribution features for electric power generation prediction

Ting-Jen Chang 1, Samantha Lee2, Jonathan Lee3, and Chi-Jie Lu1,4,5
1Graduate Institute of Business Administration, Fu Jen Catholic University, New Taipei City 242062, Taiwan
2Department of Industrial Engineering and Operations Research, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720 USA
3Division of Computing, Data Science, and Society, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720 USA
4Department of Information Management, Fu Jen Catholic University, New Taipei City 242062, Taiwan.
5 Artificial Intelligence Development Center, Fu Jen Catholic University, New Taipei City 242062, Taiwan.

Corresponding author: Chi-Jie Lu (e-mail: 059099@mail.fju.edu.tw).

This work was supported in part by the Ministry of Science and Technology, Taiwan under Grant 105-2221-E-231-005-MY2, and in part by the Fu Jen Catholic University under Grant A0110181.

ABSTRACT Developing an effective interval-valued time series (ITS) forecasting scheme for electric power generation is an important issue for energy operators and governments when making energy strategic decisions. The existing studies for ITS forecasting only consider basic descriptive information such as center, radius, upper and lower bounds, and overlooks the distribution information within the data interval. In this study, an interval-valued time series forecasting scheme based on probability distribution information features of interval-valued data with machine learning algorithms is proposed to enhance electric power generation forecasting. In the proposed scheme, the central tendency features and dispersion features from the interval-valued data are designed as integrated features sets (IFS) and used as predictor variables. Three methods including super vector regression and extreme learning machine and multivariate adaptive regression splines based on the IFS are utilized to develop ITS forecasting models. The daily time series of the metered generation from the Australian Energy Market Operator is used to illustrate the proposed scheme. Empirical results show that the proposed ITS forecasting schemes with IFS outperform the eight benchmark models and thus validate that the proposed scheme is an effective alternative for interval-valued electric power generation forecasting.

INDEX TERMS Interval-valued forecasting, power generation prediction, machine learning, probability distribution, feature engineering.

I. INTRODUCTION

As the effects of climate change intensify, governments begin adopting various methods of electric power generation, which led to an overall increase in the percentage of green power usage. However, since the efficiency of green power is largely impacted by the environment, the stability of the overall power supply remains uncertain [1],[2],[3]. Due to the essential role that electric power supply plays in a nations’ economic development, developing an effective power supply forecasting schemes has been one of the top priorities [4],[5].

Methods of forecasting electric or green power supply can be found in many pieces of literature. These methods can be divided into points forecasting, interval forecasting, and interval-valued time series (ITS) forecasting schemes. Base on the points forecasting scheme, Sobri et al. [6] discussed how time series statistical methods, physical methods, and ensemble methods outperformed traditional methods in solar PV generation forecasting. Das et al. [7] analyzed different PV power forecasting models from past publications. Li et al. [8] used a hybrid model of Long Short-Term Memory (LSTM) and Wavelet packet decomposition to improve photovoltaic (PV) power generation forecasting. Abedinia et al. [9] enhanced the empirical mode decomposition used in a hybrid model of bagging neural network and K-means to forecast wind power. Jung et al. [10] improved solar PV generation forecasting using LSTM for temporal pattern recognition.
Pazikadin et al. [11] documented the applications of Artificial Neural Network (ANN) on solar power generation forecasting. Sharadga et al. [12] tested the performance of different statistical and artificial intelligence methods on forecasting PV output. Although point forecasting is one of the most commonly adopted schemes in publications related to green power supply or power generation [13],[14], it fails to exhibit the range of variability in time series data [13],[15],[16].

Base on the interval forecasting scheme, which utilizes confidence intervals of forecasting error or direct ranges to range forecasting to predict the range of future values [13],[17],[18]. For example, Da Morita et al. [19] proposed a forecasting scheme for power load monitoring by incorporating grey prediction with confidence intervals. Da Silva and Moulin [20] proposed a hybrid scheme combining single-valued predictions of ANNs and confidence intervals estimated using multiple regression. Jiang and Li [21] designed a composite forecasting framework that combined feature extraction and fuzzy set theory selection technique to estimate the prediction intervals of wind speed time series through the machine learning method embedding a multi-objective salp swarm algorithm. Khasheia et al. [22] compared the four different confidence interval ARIMA-based time series methods in financial markets forecasting. Kim et al. [15] utilize AR, BOOT (AR model using the bias-corrected bootstrap), SARIMA, ETS (state-space exponential smoothing model), and ST (Harvey’s structural time series models) to estimate confidence intervals of the passenger arrivals in Hong Kong and Macao. Since interval forecasting often requires making assumptions or statistical inferences about the distributions of forecasting error and other prior information [13], the interval-valued forecasting scheme, which aims to deal with interval-valued time series data, rose to popularity.

Interval-valued time series (ITS) is a type of data format composed of maximum and minimum values of the interval data as upper and lower bound [23]. Many works of literature have been used the ITS forecasting model to deal with various time series data forecasting tasks, such as stock prices, exchange rate, electricity prices, electric power demand, etc. [13]. Besides, ITS forecasting has been widely used in the power industry. Roque et al. [24] constructed the iMLP model by applying it to predict daily electricity prices intervals as a function of the generation mix in the Spanish electricity market. García-Ascanio & Maté [25] forecasted Spain’s monthly electric power demand to compare the performance of a new approach of vector autoregressive forecasting models applied to interval time series with that of iMLP, or multi-layer perceptron model adapted to interval data. Results of their study reveal that iMLP shows produce more accurate forecasts for daily electricity price intervals. Rana et al. [26] use the lower and upper bound of Australian photovoltaic data to devise a new approach, named 2D-interval forecasts. It models SVR by previous historical solar power and meteorological data, and the result provides more accurate predictions than several baselines and other methods. Gligorić et al. [27] incorporate the uncertainties in the electricity price forecasting model by expressing electricity prices as interval values. The proposed hybrid model utilizes fuzzy C-mean clustering and the ITS autoregressive process to predict the short-term electricity price. Pekaslan, et al. [28] used a fuzzy logic system to perform not only a numeric forecast of the power generation but also prediction by uncertainty intervals. The result indicates that a complete fuzzy system output can provide

| Forecasting Schemes and Ref. | Characteristics | Limitation |
|------------------------------|-----------------|------------|
| Points Forecasting Scheme [6],[7],[8],[9],[10],[11],[12], [14] | It is a simple method to represent the expected value of the outcome, which depends on how likely they are to happen to weigh different possible outcomes by assumption. | Because it estimates the unknown true future value by a single number, it cannot provide information as to the degree of uncertainty associated with the forecast. |
| Interval Forecasting Scheme [19],[20],[21],[15], [22] | It provides a margin of error around the point to indicate a range of possible future outcomes with a prescribed level of confidence, and it is also extended to predict the range of future values. | It assumes that observations and estimations are usually incomplete or uncertain, meanwhile ignoring the sampling variability related to parameter estimation, so it underestimates the degree of future uncertainty. |
| ITS Forecasting Scheme [23],[13], [23],[24], [25],[26], [27],[28] | It is a data-driven method, which is allowed new variable types to be introduced. The properties of ITS are described by these interval-valued variables. | Most research considers intuitive variables like the upper, lower, width, radius, and middle points only. |
relevant more information in terms of the reliability of the forecast using wind power ITS datasets.

As above-mentioned literature, most researches apply upper-lower bound (UL) or center-radius (CR) schemes to construct the interval-valued data for modeling. However, in the interval-valued data constructed by the UL-based scheme, only the information of extreme value is considered, while the structure and distribution information of intra-interval data are often overlooked. Likewise, the CR-based scheme assumes the data distribution within the interval is symmetrical to the center point, and that the radius from the center point to the upper and lower bound is consistent. It also does not consider the distribution or dispersion information in the interval. In reality, there is much information between the upper and lower bounds of interval-valued data that can be used to generate useful features for building an effective ITS forecasting model. However, existing studies for ITS forecasting only considered the upper and lower bound, or center and radius information in the interval, and pay little attention to the distribution information, such as central tendency and dispersion tendency features, within an interval-valued data.

In this study, an interval-valued time series forecasting scheme with probability distribution information features (PF) and machine learning (ML) algorithms are proposed for electric power generation forecasting. In the proposed scheme, the UL-based and CR-based schemes are used to construct daily interval-valued electric power generation data. The central tendency features and dispersion features from the intra-interval data are designed and generated as predictor variables. Three ML methods including supper vector regression (SVR) and extreme learning machine (ELM) and multivariate adaptive regression splines (MARS) based on the designed predictor variables are utilized to develop forecasting models.

The three forecasting algorithms used represent different perspectives for modeling. SVR is a powerful regression method based on statistical learning principal [29],[30]; ELM is one algorithm of the feedforward neural networks for classification, regression, clustering, sparse approximation, compression, and feature learning [31]; MARS is a nonlinear multivariate algorithm by producing a piecewise linear model automatically [32]. To compare with popular statistical methods for electric power generation forecasting, random walk (RW), exponential smoothing state space model (ETS), and autoregressive integrated moving average (ARIMA) are also used in this study. It is common to find these methods applied in a research project on electric power generation forecasting. Examples include using the SVR-based model to forecast photovoltaic power generation [34],[35],[36]; an ELM-based model to predict power generation [37],[38]; a MARS-based model to predict power generation [39],[40],[41]; and ARIMA model for electric power generation forecasting [42],[43]. Therefore, SVR, ELM, and MARS indeed provide mainstream and different frameworks to our research.

The rest of this paper is organized into 4 sections. Section 2 gives a brief introduction to SVR, ELM, and MARS. In section 3, we present how the research process constructs the proposed forecasting models. The experimental results and related discussions are presented in Section 4. In the end, the paper gives a conclusion in section 5.

II. Methods

A. SUPPORT VECTOR REGRESSION

Support vector regression (SVR) is a supervised ML algorithm method for solving nonlinear regression estimation problems by transforming the nonlinear input area to an area with high-dimensional properties to find a hyper-plane through nonlinear mapping that helps predict the target value. The SVR model can be constructed as follows [44],[45]:

\[
f(x) = (z \cdot \phi(x)) + b
\]

where \(z\) is the weight vector, \(b\) is the bias term, and \(\phi\) is the representation of the nonlinear mapping function that maps \(x\) into a higher dimensional feature space.

In addition, SVR introduces an \(\epsilon\)-sensitive region around the function and seeks to minimize the corresponding \(\epsilon\)-sensitive convex loss function [44]. The loss function is defined as [44],[45]:

\[
L(f(x) - y) = \begin{cases} |f(x) - y| - \epsilon & \text{if } |f(x) - y| > \epsilon \\ 0 & \text{otherwise} \end{cases}
\]

where \(y\) is the desired output and epsilon is the \(\epsilon\)-sensitivity region. As shown by the definition of the loss function, if the predicted value falls within the region, the loss is zero. Otherwise, the loss is the magnitude of the difference between the predicted value and the value of \(\epsilon\). Hence, we can formulate SVR as a minimization quadratic programming problem as below [45],[46]:

\[
\begin{align*}
\text{Min} & : \frac{1}{2}z^Tz + C \sum_{i}(\xi_i^+ + \xi_i^-) \\
\text{Subject to} & : \\
& y_i - z^T\hat{x}_i - b \leq \epsilon + \xi_i^- \\
& z^T\hat{x}_i + b - y_i \leq \epsilon + \xi_i^+ \\
& \xi_i^+, \xi_i^- \geq 0
\end{align*}
\]

where \(l = 1,..., n\) represents the number of data in the training set; \(\xi_i, \xi_i^+\) represent the distance from the actual values to the boundaries of the E-sensitive region; \((\xi_i^+ + \xi_i^-)\) is the empirical risk; \(1/2z^Tz\) is the regularization term preventing overfitting; \(C\) is the regularization constant representing the trade-off between the regularization term and empirical risk.

We can formulate the dual program of the above problem by using the Lagrange multipliers and selecting appropriate \(C, \epsilon\) and kernel function that meets Mercer’s condition. This yields the general form of the SVR-based regression function [44],[45]:

\[
f(x, z) = f(x, \alpha, \alpha^*) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\phi(x_i, x_j) + b
\]
where $\alpha_i$ and $\alpha_i^*$ are Lagrangian multipliers and satisfy the equality $\alpha_i \ast \alpha_i^* = 0$.

**B. EXTREME LEARNING MACHINE**

Extreme learning machine (ELM) is a single hidden layer feedforward neural network (SLFN) in which the input weights are randomly initialized and fixed to analytically determine the weights of the output using an activation function [46]. ELM yields promising performance due to its efficient learning speed and fast convergence, as it avoids problems commonly faced by traditional gradient-based algorithms such as converging at local extrema, over-tuning, inefficient learning speed and fast convergence, as it avoids

The MARS algorithm is given as [47],[48],[49]:

$$H_{km}(x_{v(k,m)}) = \prod_{k=1}^{K_m}[s_{k,m}(x_{v(k,m)} - t_{k,m})]$$

where $s_{k,m}$ takes values $\pm 1$, $x_{v(k,m)}$ labels the predictor variables, $t_{k,m}$ indicates the knot locations, and $K_m$ denotes the number of “splits” that generate the $m$-th basis function. The process of computing the optimal MARS model consists of two phases: the forward pass and the backward pass. The forward pass involves creating a collection of basic functions and add them iteratively in pairs to build an overfitting model. In particular, the algorithm finds the pair producing the greatest improvement in the model error at each step. On the other hand, the backward pass serves to enhance the model’s generalization ability through pruning. Specifically, basis functions are eliminated in the order of least contributions based on the generalized cross-validation (GCV) criterion, which is defined as [47],[48],[49]:

$$GCV(M) = \frac{1}{N} \sum_{i=1}^{N} \frac{|y_i - f_M(x_i)|^2}{1 + \frac{C(M)^2}{N}}$$

where $N$ is the number of observations and $M$ is the number of basis functions. The denominator is a complexity function with $C(M)$ being the cost-penalty measure that penalizes model complexity and avoids overfitting; the numerator measures the lack of fit on $f_M(x_i)$, the model of the M basis functions. The pruning process stops when it finds the best sub-model and all the remaining basis functions satisfy the pre-determined requirements.

**III. Proposed ITS Forecasting Scheme with PF and ML**

Figure 1 shows the flowchart of the proposed ITS forecasting scheme with PF and ML. As the figure depicts, the daily electric power generation intraday raw data are used to construct daily interval-valued electric power generation data by using UL and CR-based schemes, respectively. The lower bound, upper bound, center, and radius are referred to as basic features (BF) in this study. Additionally, the central tendency features including mean and median, and the dispersion features including interquartile range, standard deviation, skewness, and kurtosis can be calculated from intraday electric power generation data and named as probability distribution information features (PF) in this study. The PF will be used to combine with each BF to generate four designed predictor variable sets for modeling. In other words, the upper bound feature combined with PF forms an integrated feature set and serves as the input of an ML model to predict the upper bound; in the same manner, the lower bound, center and radius features are, respectively, combined with PF to construct integrated feature sets ML model to predict lower bound, center and radius individually. The predicted center and radius will be converted into lower bound and upper bound as the final results of the CR-based scheme. Lastly, the metrics of the
mean squared error of interval (MSEI) and mean relative interval error (MRIE) are used to evaluate the performance of ML models. The detailed depiction of the proposed scheme is shown below.

**STEP 1 COLLECT INTRADAY RAW DATA**
Suppose an intraday time series data $\mathbf{p}$ of length $n$ are collected, $\mathbf{p} = (p_1, p_2, \cdots, p_n)$. Figure 2 shows intraday electric power generation data as an example. Below is an intraday time series of 2 days containing 96 data points, which consist of data values collected every 30 minutes from 00:00 to 24:00 every day. A part of the data points is expressed in Table 2.

**STEP 2 CONSTRUCT DAILY INTERVAL-VALUED DATA**
Considering the time length $t$ as the time period/interval, which can be every half hour, hour, day, week, or month, the time series $\mathbf{p}$ can be converted into a $d$ interval-valued data, and get a matrix $\mathbf{X}_d$ of size $d \times t$ as follows,
\[ X_D = \begin{bmatrix} p_1 & \ldots & p_t & \ldots & p_{d-1} & p_d \\ p_t+1 & \ldots & p_{d+t} & \ldots & p_{2d-1} & p_{2d} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ p_{(d-1)t+1} & \ldots & p_{dt} & \ldots & p_{d^2-1} & p_{d^2} \end{bmatrix} = \begin{bmatrix} x_1 \ldots x_{1t} \\ x_2 \ldots x_{2t} \\ \vdots \ddots \vdots \\ x_d \ldots x_{dt} \end{bmatrix} \]

where \( d = \frac{t}{r} \)

For the data in Figure 2 and Table 2, the daily intraday time series \( p = \{821, 831, \ldots, 863, 998\} \) can be transformed as a matrix \( X_D \) by using \( t=48 \) (48 data points in one day) and listed in Table 3.

**STEP 3 EXTRACT BASIC FEATURES**

For BF extraction, the maximum value \( x^U_i \) and minimum value \( x^L_i \) in the intraday data denote the upper bound and lower bound used to construct a matrix \( X_H \) of UL-based scheme. The process is expressed as,

\[ x^U_i = Max(x_i), x^L_i = Min(x_i), \]
\[ x^L_i < x^U_i, i = 1,2,\ldots,d \]  

\[ X_H = \begin{bmatrix} x^U_1 & x^U_2 & \ldots & x^U_d \\ x^L_1 & x^L_2 & \ldots & x^L_d \end{bmatrix} \]  

Then, a matrix \( X_R \) for the CR-based scheme is created using the center \( x^C_i \) and radius \( x^R_i \), where \( x^C_i \) and \( x^R_i \) can be calculated from \( x^U_i \) and \( x^L_i \),

\[ x^C_i = \frac{x^U_i + x^L_i}{2}, x^R_i = \frac{x^U_i - x^L_i}{2}, i = 1,2,\ldots,d \]  

\[ X_R = \begin{bmatrix} x^C_1 & \ldots & x^C_d \\ x^R_1 & \ldots & x^R_d \end{bmatrix} \]  

Note that the upper and lower bounds of the CR-based scheme can also be calculated using center and radius.

\[ x^U_i = x^C_i + x^R_i, x^L_i = x^C_i - x^R_i, i = 1,2,\ldots,d \]  

Using Table 3 as an example, a matrix \( X_H \) and \( X_R \) can be obtained by using (12) to (15),

\[ X_H = \begin{bmatrix} x^U_1 & x^U_2 \\ x^L_1 & x^L_2 \end{bmatrix} = \begin{bmatrix} 1,284 & 1,368 \\ 805 & 819 \end{bmatrix} \]

\[ X_R = \begin{bmatrix} x^C_1 & x^C_2 \\ x^R_1 & x^R_2 \end{bmatrix} = \begin{bmatrix} 1,044.5 & 1,093.5 \\ 239.5 & 274.5 \end{bmatrix} \]

Figure 3 shows the interval structure of UL-based and CR-based schemes. \( x^C \) is between \( x^U \) and \( x^L \), where \( x^R \) is the distance between \( x^C \) and both \( x^U \) and \( x^L \). The \( x^U \) and \( x^L \) present information relevant to the range of ITS while \( x^C \) and \( x^R \) provide the range and center information of ITS.
FIGURE 3. Interval structure based on upper bound, lower bound, center, and radius

STEP 4 EXTRACT PROBABILITY DISTRIBUTION FEATURES

For PF, the central tendency and dispersion information of ITS are extracted. The mean ($x^M$), standard deviation ($x^{SD}$), median ($x^{Med}$), interquartile range ($x^I$), skewness ($x^{Sk}$), and kurtosis ($x^{Ku}$) of ITS are computed and used to construct a PF matrix $X_F$, shown in the following,

$$X_F = \begin{bmatrix} x^M & x^M & \cdots & x^M \\ x^{SD} & x^{SD} & \cdots & x^{SD} \\ x^{Med} & x^{Med} & \cdots & x^{Med} \\ x^I & x^I & \cdots & x^I \\ x^{Sk} & x^{Sk} & \cdots & x^{Sk} \\ x^{Ku} & x^{Ku} & \cdots & x^{Ku} \end{bmatrix}$$

(17)

where $x^M_i = \frac{1}{t} \sum_{j=1}^{t} x^M_{ij}$.

$$x^{SD}_i = \sqrt{\frac{1}{t} \sum_{j=1}^{t} (x^M_{ij} - x^M)^2}$$

if $i$ is odd number

$$x^{Med}_i = \begin{cases} \frac{x^M_{i+1}}{2} & \text{if } i \text{ is even number} \\ \frac{1}{2} (x^M_{i} + x^M_{i+1}) & \text{if } i \text{ is odd number} \end{cases}$$

$$x^{Q1}_i = \begin{cases} \frac{x^M_{i+1}}{4} & \text{if } i \text{ is even number} \\ \frac{1}{2} (x^M_{i} + x^M_{i+1}) & \text{if } i \text{ is odd number} \end{cases}$$

$$x^{Q3}_i = \begin{cases} \frac{x^M_{i+1}}{4} & \text{if } i \text{ is even number} \\ \frac{1}{2} (x^M_{i} + x^M_{i+1}) & \text{if } i \text{ is odd number} \end{cases}$$

$$x^I_i = x^{Q3}_i - x^{Q1}_i$$

$$x^{Sk}_i = 3 \frac{(x^M - x^{Med})}{x^{SD}}$$

$$x^{Ku}_i = \frac{1}{t} \sum_{j=1}^{t} (x^M_{ij} - x^M)^4 \left( \frac{1}{t} \sum_{j=1}^{t} (x^M_{ij} - x^M)^2 \right)^{-3}$$

$i = 1,2, \cdots, d \quad j = 1,2, \cdots, t.$

There are two possible kinds of distributions in an ITS. The first is shown in Figure 4(a). The distribution is symmetrical with mean equal to median and $x^{Sk} = 0$. If $x^{Ku} \approx 0$ is also true, the intraday data is following the Gaussian distribution. It is a hypothetical condition constructed based on traditional statistics by UL- and CR-based schemes, but it is hardly practical in real-world situations. Most data distributions are asymmetrical, as shown in Figure 4(b). Moreover, many possible shapes of distribution exist in the interval-valued data. When the mean is smaller than the median, the distribution is skewed left and the skewness is smaller than 0. On the other hand, when the distribution is skewed to the right, the skewness is higher than 0. Thus, the PF can be helpful in describing the information contained in the data distribution within ITS.

(a). Symmetrical distribution

(b). Skewed distribution

FIGURE 4. Interval structure based on upper bound, lower bound, center, and radius
STEP 5 CONSTRUCT INTEGRATED FEATURES SETS FOR MODELING

After obtaining the BF and PF, we combined each single BF with the whole PF respectively to construct an integrated feature set (IFS) for modeling. In other words, four BFs including $\mathbf{x}_1^L$, $\mathbf{x}_1^C$, $\mathbf{x}_1^R$, and $\mathbf{x}_2^L$ are individually added to the PF matrix $\mathbf{X}_F$ to generate four different IFS. It can be expressed as,

$$\mathbf{X}^{PF(\beta)} = \begin{bmatrix} x_1^\beta & x_2^\beta & \ldots & x_d^\beta \\ x_1^M & x_2^M & \ldots & x_d^M \\ x_1^{Med} & x_2^{Med} & \ldots & x_d^{Med} \\ x_1^{Sk} & x_2^{Sk} & \ldots & x_d^{Sk} \\ x_1^{SD} & x_2^{SD} & \ldots & x_d^{SD} \\ x_1^{Ku} & x_2^{Ku} & \ldots & x_d^{Ku} \end{bmatrix}$$

where $\beta = \{U, L, C, R\}$. 

In this study, the proposed scheme considers one to five time-lags information to predict lower, upper bound, center, and radius on the day $s$ based on $\mathbf{X}^{PF(\beta)}$ and it can be expressed as,

$$\hat{x}_s^\beta = f(x_{s-1}^\beta, x_{s-1}^M, x_{s-1}^{Med}, x_{s-1}^{Sk}, x_{s-1}^{SD}, x_{s-1}^{Ku})$$

where $\beta = \{U, L, C, R\}$. 

To evaluate the performance of the proposed scheme, a model containing single BF to predict the lower bound, upper bound, center, and radius, is utilized as the baseline and can be expressed as,

$$\hat{x}_s^\beta = f(x_{s-1}^\beta), i = 1, 2, \ldots, 5, \text{ and } \beta = \{U, L, C, R\}$$

Note that, $\hat{x}_s^C$ and $\hat{x}_s^R$ in (19) and (20) are the predicted values of center and radius, respectively. The predicted lower bound $\hat{x}_s^L$ and upper bound $\hat{x}_s^U$ are obtained using (20).

$$\hat{x}_s^L = \hat{x}_s^C + \hat{x}_s^R, \hat{x}_s^U = \hat{x}_s^C - \hat{x}_s^R, i = 1, 2, \ldots, d$$

STEP 6 CONSTRUCT SVR, ELM, AND MARS MODELS

The four IFSs are used as input variables for constructing SVR, ELM, and MARS models. The UL-PF-SVR, UL-PF-ELM, and UL-PF-MARS with UL-based scheme are used to predict the upper bound and lower bound directly. The CR-PF-SVR, CR-PF-ELM, and CR-PF-MARS apply a CR-based scheme to predict the upper bound and lower bound. The UL-SVR, UL-ELM, UL-MARS, CR-SVR, CR-ELM, and CR-MARS models using a single BF as input variable serve as comparison models to help evaluate the performance of the proposed models.

Because RW, ETS, and ARIMA are common univariate models in time series forecasting, we also consider constructing RW, ETS, and ARIMA by UL- and CR-based scheme. So, the forecasting performance of UL-RW, CR-RW, UL-ETS, CR-ETS, UL-ARIMA, and CR-ARIMA is compared with that of our proposed scheme.

STEP 7 EVALUATE PERFORMANCE

Hsu & Wu [50] proposed two promising metrics, namely the mean squared error of interval (MSEI) and mean relative interval error (MRIE) to evaluate forecasting performance for interval-valued data. Since MSEI and MRIE are calculated by center and radius, if the predicted values are pure lower bound and upper bound, they need to be expressed in terms of center and radius using (16). The definitions of the interval-valued metrics are summarized in Table 4. The smaller the values are, the greater the accuracy.

| METRICS | CALCULATION* |
|---------|--------------|
| MSEI    | $\frac{1}{m} \sum_{j=1}^{m} (C_j - \hat{C}_j)^2 + \frac{1}{2} \sum_{j}^{m} (R_j - \hat{R}_j)^2$ |
| MRIE    | $\frac{1}{m} \sum_{j}^{m} |C_j - \hat{C}_j|$ |

* Note that $C$ and $\hat{C}$ represent the predicted and actual value of the center of an interval-valued data; $R$ and $\hat{R}$ represent the predicted and actual value of the radius of an interval-valued data; $m$ is the data.

IV. EMPIRICAL STUDY

Australia is facing the change of evolving power supply mix, aging electrical infrastructure, weather, and increasing interdependencies on electricity markets. The Australian Energy Market Operator (AEMO) is responsible for managing the electrical systems and markets across Australia, helping the government to maintain secure electricity systems, and leading the design of Australia’s future energy system. To evaluate the performance of the proposed forecasting model with IFS of ITS, we pick up the time series of the metered generation from AEMO’s database https://data.wa.aemo.com.au/#load-summary. The metered generation forecasting is important for energy trading through a coordinated dispatch process. The time series of metered generation is collected by a 30-minute period commencing on the hour.

The time series in this study begins from 2017/2/1 00:00 to 2019/10/31 23:30 shown in Figure 5. Excluding the annual inspection period, there are 45,264 data points in total. As there are 48 intraday data points every day, using the method mentioned in Section 3 gives 943 data points for the lower and upper bound, derived from the minimum and maximum intraday values as shown in Figure 6. The first 80% of ITS, or 755 data points, are used as the training data set while the remaining 20%, containing 188 data points, are used as the out-of-sample testing data set.

Using the modeling process mentioned in Section 3, fourteen prediction models are constructed and their prediction performances are listed in Table 5. From Table 5, we can see that the MSEI and MRIE of CR-PF-MARS are not only lower than the other five ML models using IFS but also superior to the other twelve comparison models not using IFS. Thus, we can infer that CR-PF-MARS is the best model of interval-valued time series forecasting for electric power generation.
In Figure 7, the predicted results of ML methods with IFS are plotted. The trend of the prediction result is close to the actual value. It means IFS can provide useful information for ML models to capture fluctuations of electric power generation.

To compare the performance between ML with and without IFS, we calculate and summarize their difference in Table 6. Taking the CR-PF-MARS and CR-MARS as examples, the MSEI of CR-PF-MARS is 3,440.23, and CR-MARS is 3,953.25. The delta of CR-PF-MARS compared to CR-MARS is negative 12.98%, which means that the MSEI of CR-PF-MARS is lower than CR-MARS 12.98%. Using the same approach, the MRIE of CR-PF-MARS is 7.5% lower than CR-MARS. We can find that the six models we proposed are more accurate than their single BF counterparts. Thus, we can infer that IFS provides more information and improves the performance of forecasting models.

**TABLE V**

| Methods     | MSEI     | MRIE |
|-------------|----------|------|
| UL-PF-SVR   | 3,731.35 | 0.078|
| UL-PF-ELM   | 3,649.34 | 0.076|
| UL-PF-MARS  | 3,543.56 | 0.077|
| CR-PF-SVR   | 3,643.35 | 0.077|

**FIGURE 5.** Electric power generation of the 30-minute period commencing on the hour in Australia from 2017/2/1 00:00 to 2019/10/31 23:30.

**FIGURE 6.** The daily electric power generation of upper and lower bound of Figure 5.
The predicted results of machine learning methods with IFS.

FIGURE 7. The predicted results of machine learning methods with IFS.

TABLE VI
THE IMPROVED RATE OF PREDICTION PERFORMANCE THROUGH ML WITH AND WITHOUT IFS

| Scheme          | MSEI   | Improve Rate | MRIE  | Improve Rate |
|-----------------|--------|--------------|-------|--------------|
| UL-PF-SVR       | 3,731.35 | -2.33%       | 0.078 | -2.50%       |
| UL-SVR          | 3,820.29 | 0.080        |       |              |
| UL-PF-ELM       | 3,649.34 | -6.38%       | 0.076 | -5.00%       |
| UL-ELM          | 3,897.90 | 0.080        |       |              |
| UL-PF-MARS      | 3,543.56 | -8.42%       | 0.077 | -2.53%       |
| UL-MARS         | 3,869.23 | 0.079        |       |              |
| CR-PF-SVR       | 3,643.35 | 0.077        | 0.079 | -2.53%       |
| CR-SVR          | 3,813.52 | -4.46%       | 0.079 | -2.53%       |
| CR-PF-ELM       | 3,456.19 | -9.35%       | 0.075 | -5.06%       |
| CR-ELM          | 3,812.57 | 0.079        |       |              |
To evaluate the robustness of the proposed scheme, different relative ratios of training and testing sets are used. The relative ratios for the testing plan are calculated from the size of the training dataset to the size of the complete dataset. Four relative ratios are considered: 70%, 75%, 80%, and 85%.

The electric power generation prediction results for the six proposed methods and eight comparison methods investigated herein are gathered in Table 7 in terms of the MRIE metric. Table 7 shows that the CR-PF-MARS still presents the best prediction performance under different ratios of training and testing sets. It indicates that the CR-PF-MARS methodology certainly provides higher accuracy for electric power generation prediction.

![Figure 8. The predicted results of machine learning methods with IFS.](image)

### Table VII

| Methods       | 70% MRIE | 75% MRIE | 80% MRIE | 85% MRIE |
|---------------|----------|----------|----------|----------|
| UL-PF-SVR     | 0.092    | 0.083    | 0.078    | 0.079    |
| UL-PF-ELM     | 0.090    | 0.081    | 0.076    | 0.078    |
| UL-PF-MARS    | 0.093    | 0.082    | 0.077    | 0.081    |
| CR-PF-SVR     | 0.090    | 0.082    | 0.077    | 0.078    |
| CR-PF-ELM     | 0.088    | 0.080    | 0.075    | 0.078    |
| CR-PF-MARS    | **0.088**| **0.080**| **0.074**| **0.077**|
| UL-RW         | 0.143    | 0.187    | 0.235    | 0.175    |
| UL-ETS        | 0.144    | 0.174    | 0.221    | 0.173    |
| UL-ARIMA      | 0.164    | 0.149    | 0.152    | 0.139    |
| UL-SVR        | 0.095    | 0.084    | 0.080    | 0.083    |
| UL-ELM        | 0.097    | 0.083    | 0.080    | 0.081    |
| UL-MARS       | 0.093    | 0.083    | 0.079    | 0.082    |
| CR-RW         | 0.143    | 0.187    | 0.235    | 0.175    |
| CR-ETS        | 0.154    | 0.167    | 0.225    | 0.168    |
| CR-ARIMA      | 0.165    | 0.144    | 0.187    | 0.142    |
| CR-SVR        | 0.093    | 0.082    | 0.079    | 0.081    |
| CR-ELM        | 0.092    | 0.081    | 0.079    | 0.082    |
| CR-MARS       | 0.092    | 0.082    | 0.080    | 0.081    |

Figure 8 shows the performance comparison between ML with and without IFS under the different relative ratios of training and testing sets. From this Figure, it can be observed that the proposed scheme demonstrates robustness as it reduces the forecasting errors even when different sizes of the training set are used. Taking the CR-PF-MARS vs. CR-MARS as an example, it can be observed that the performance of CR-PF-MARS on 70% ratio is 4.35% more accurate than CR-MARS as expressed in the MRIE metric. Even with 75%, 80%, and 85% training set, CR-PF-MARS continues to outperform CR-MARS in terms of MRIE. The other five ML-IFS models also show similar results. Their performances are also more accurate than the other eight comparison models. Therefore, IFS can indeed improve the predictive performance of ML models.

### VI. Conclusion

Electric power generation is not easy to predict because of evolving power supply mix is changing, as driven by weather, power generation technologies, and increasing interdependencies among green power, thermal power, nuclear power. This study proposed an effective scheme based on the designed integrated feature set to improve the ITS forecasting performance using SVR, ELM, and MARS models. The experimental results show that ML with IFS produces a lower prediction error and outperforms ML without IFS. The proposed CR-PF-MARS model can present a promising performance and is a good alternative for electric power generation forecasting.

### Reference

[1] Semero, Y. K., Zhang, J., & Zheng, D., “PV power forecasting using an integrated GA-PSO-ANFIS approach and Gaussian process regression based feature selection strategy”, CSEE Journal of Power and Energy Systems, vol. 4, no. 2, pp. 210–218, June 2018.

[2] Zhou, H., Zhang, Y., Yang, L., Liu, Q., Yan, K., & Du, Y., “Short-term Photovoltaic Power Forecasting based on Long Short-Term Memory Neural Network and Attention Mechanism”, IEEE Access, vol. 7, pp. 78063-78074, 2019.

[3] Niu, D., Wang, K., Sun, L., Wu, J., & Xu, X., “Short-term photovoltaic power generation forecasting based on random forest feature selection and CEEMD: A case study”, Applied Soft Computing, vol. 93, 106389, Aug. 2020.

[4] Niu, D., Shi, H., Li, J., & Wei, Y., “Research on short-term power load time series forecasting model based on BP neural network”, 2010 2nd International Conference on Advanced Computer Control, pp. 509-512, 2010.

[5] Guo, L., Chen, J., Wu, F., & Wang, M., “An electric power generation forecasting method using support vector machine”, Systems Science & Control Engineering, Vol. 6, Issue 3, 191–199, 2018.

[6] Sobri, S., Koohi-Kamali, S., Rahim N. A., “Solar photovoltaic generation forecasting methods: A review”, Energy Conversion and Management, Vol. 156, pp. 459-497, 2018.

[7] Das, U.K., Tey, K.S., Seyedmamoudian, M., Mekhilef, S., Idris, M.Y.I., Deventer, W.V., Horan, B., Stojevska, A., “Forecasting of photovoltaic power generation and model optimization: A review”, Renewable and Sustainable Energy Reviews, Vol. 81, Part 1, pp. 912-928, Jan. 2018.

[8] Li, P., Zhou, K., Lu, X., Yang, S., “A hybrid deep learning model for short-term PV power forecasting”, Applied Energy, Vol. 259, 114216, Feb. 2020

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3142083, IEEE Access
[49] Lu, C. J., Lee, T. S., & Lian, C. M. “Sales forecasting for computer wholesalers: A comparison of multivariate adaptive regression splines and artificial neural networks”, Decision Support Systems, vol. 54, issue 1, pp. 584-596, Dec. 2012.

[50] Hsu, H.-L., & Wu, B., “Evaluating forecasting performance for interval data”, Computers & Mathematics with Applications, vol. 56, issue 9, pp. 2155–2163, Nov. 2008.

Ting-Jen Chang received the B.S. in history and M.B.A degrees from Fu Jen Catholic University, New Taipei City, Taiwan, in 1997 and 2006. He is currently pursuing the Ph.D. degree in business administration with the Fu Jen Catholic University. His research interests include time series forecasting, feature engineering, and machine learning.

Samantha Lee is currently pursuing bachelor’s degree in IEOR and EECS at the University of California, Berkeley. Her research interests include big data analytics, machine learning, and data mining.

Jonathan Lee is currently pursuing a Data Science (B.A.) and Computer Science (B.A.) double major degree in University of California, Berkeley. His current research interests include big data analysis, machine learning, and artificial intelligence related topics.

Chi-Jie Lu is currently a professor in the Graduate Institute of Business Administration & Department of Information Management at Fu Jen Catholic University, Taiwan. He got his Ph.D. in Industrial Engineering and Management from Yuan-Ze University, Taiwan. His research areas include machine learning and its applications, time series data analysis, medical/healthcare Informatics, supply chain management and quality management. He has published over 50 refereed articles in various journals, including Pattern Recognition, Decision Support Systems, Neurocomputing, International Journal of Production Economics, Engineering Applications of Artificial Intelligence, European Journal of Operational Research, Neural Computing and Applications, and Computational Statistics and Data Analysis.