

Continuous Conditional Random Field Model for Predicting the Electrical Load of a Combined Cycle Power Plant

Gilseung Ahn, Sun Hur*

Department of Industrial and Management Engineering, Hanyang University, Ansan, Korea

(Received: February 14, 2016 / Revised: May 1, 2016 / Accepted: June 5, 2016)

ABSTRACT

Existing power plants may consume significant amounts of fuel and require high operating costs, partly because of poor electrical power output estimates. This paper suggests a continuous conditional random field (C-CRF) model to predict more precisely the full-load electrical power output of a base load operated combined cycle power plant. We introduce three feature functions to model association potential and one feature function to model interaction potential. Together, these functions compose the C-CRF model, and the model is transformed into a multivariate Gaussian distribution with which the operation parameters can be modeled more efficiently. The performance of our model in estimating power output was evaluated by means of a real dataset and our model outperformed existing methods. Moreover, our model can be used to estimate confidence intervals of the predicted output and calculate several probabilities.

Keywords: Continuous Conditional Random Field, Machine Learning, Combined Cycle Power Plant, Energy Saving, Prediction

* Corresponding Author, E-mail: hursun@hanyang.ac.kr

1. INTRODUCTION

As the demand for electric power has grown rapidly during the past several decades, so has the interest in the combined cycle power plant (CCPP). This is because CCPPs are known to be very efficient and require relatively low investment costs. A CCPP is composed of a gas turbine, steam turbine, and heat recovery system generators. The two turbines are combined in one cycle and the heat or gas flow transfers the energy from one of the turbines to the other. In general, a gas turbine exhausts gas that is used to produce heat, which is used to make the steam required by the steam turbine (Niu and Liu, 2008).

Numerous control strategies have been developed to reduce CCPP operational costs, but still a more advanced

control strategy is necessary to further reduce the entire operational cost. Tüfekci (2014) suggests that it is essential for a base load power plant to predict electrical power outputs correctly in order to attain a maximum profit. Existing plants, however, consume a significant amount of fuel and have high operating expenses partly because of poor prediction of electrical power output requirements. Particularly, the reliability and sustainability of the gas turbine are highly affected by the prediction of the power generation needs.

Some studies adopting thermodynamic approaches to obtain an accurate prediction for the power generation have been done. In order to forecast the power generation accurately with these approaches, however, many assumptions, such as the existence of some empirical relationships, are necessary since they account for unpre-

dictability in their solution. Without these assumptions, any analysis of a real application calls for many nonlinear equations, whose solution is either almost impossible or requires too much computational time and effort, and sometimes the result is still unsatisfactory and unreliable (Kesgin and Heperkan, 2005).

Several studies employing machine learning methods that enable electrical power prediction as an alternative analysis to overcome these difficulties have been conducted. In (Kesgin and Heperkan, 2005), an artificial neural network and fuzzy logic are utilized to analyze various thermodynamic systems, including a CCPP. In Fan *et al.* (2016), the authors point out that electric load forecasting is very important for power utility and they present a support vector regression (SVR) model blended with differential empirical mode decomposition (DEMD) and auto regression to forecast electric load. Yadav and Srinivasan (2011) proposes a method for short-term load forecasting, which is based on a smooth transition autoregressive (STAR) model. In Kaya *et al.* (2012), the power of combined gas and steam turbines is predicted by means of a k-nearest neighbor smoother, multivariate linear regression, artificial neural network, and some other methods. Recently, Tüfekci (2014) has dealt with several machine learning regression methods for the predictive analysis of a thermodynamic system, which is a combined cycle power plant with one steam turbine and two heat recovery systems. In Clifton *et al.* (2013), the authors perform a study to predict a wind turbine's power output by means of a regression tree. In Prokop *et al.* (2013), an application of evolutionary fuzzy rules is presented to model and predict the power output of a real-world photovoltaic power plant (PVPP). In Yu and Xu (2014), a combinational approach based on improved back propagation (BP) neural network for short-term gas load forecasting is proposed and the genetic algorithm is employed to optimize the network. In Al-Rashidi and El-Naggar (2010), the authors suggest a

novel method for annual peak load forecasting in electrical power system, which employs a particle swarm optimization to find the optimal model parameters of the model. Xie and Hong (2015) presents an integrated probabilistic electric load forecasting solution, which consists of three components: pre-processing, forecasting, and post-processing. In the forecasting component, time series modeling and neural networks are employed to forecast the electric load. With the increase of the researches employing machine learning methods to predict electrical power, several researches compare machine learning methods.

Because the gas turbine power output mostly depends on the ambient parameters such as ambient temperature, atmospheric pressure, and relative humidity, these are the input variables of our model. Moreover, exhaust steam pressure should be included in the input variables since steam turbine power output has a direct relationship with the exhaust vacuum level. Figure 1 shows the configuration of the gas turbine. Since various relationships exist among dataset variables for the prediction of electrical power output, a model that can deal well with complicated structure, including a wide variety of arbitrary and non-independent input features is desirable to obtain an accurate prediction. To the best of our knowledge, however, there is no previous research result considering various relationships among variables in the described data set. In order to consider the various relationships among the variables, a continuous conditional random field (C-CRF) is employed in this paper to predict the full load electrical power output of a base load operated combined cycle power plant.

The C-CRF model is appropriate for this regression problem for two reasons. First, the C-CRF model can accommodate many input variables and represent complex dependence relations among them into the mathematical description. Second, the C-CRF model can provide not only point estimation, but also an interval esti-

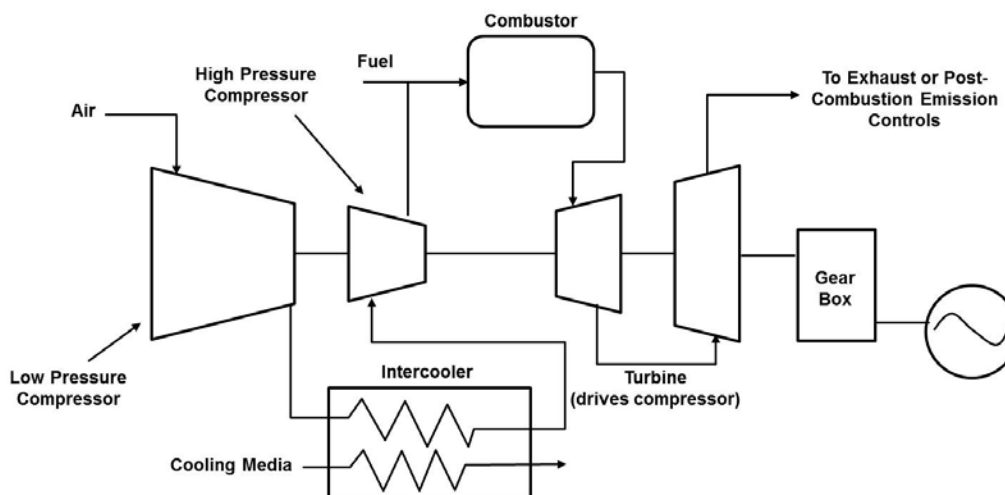


Figure 1. Configuration of the gas turbine (<http://www.understandingchp.com>).

mation of the predicted value by considering similarity among the outputs. CRF was originally developed for classification of sequential data (Lafferty *et al.*, 2011), and has been adapted for many applications in various areas, including computer vision (Kumar and Hebert, 2003). Recently, CRF has been extended to regression by allowing the target variable to be continuous (C-CRF) and applied to regression on spatial-temporal data (Liu *et al.*, 2004).

In this study, we construct an association potential, which is part of the C-CRF model, using machine learning methods: artificial neural network (ANN), regression tree (RT), and multiple linear regression (MLR). These methods have shown high predictive validity in electrical power output prediction problems. Euclidean similarity is introduced for interaction potential, which is another part of C-CRF. The model is trained with the help of a multivariate Gaussian distribution, and then is evaluated with a real data set. The prediction accuracies of the suggested model are compared with those in (Tüfekci, 2014). Additionally, some application examples are provided in order to show the effectiveness of the model.

The remainder of this paper is organized as follows. We introduce the C-CRF in Section 2 and show how to apply it to predict the power of a combined cycle power plant in Section 3. In Section 4, we evaluate the suggested model with a real data set, and compare the model with those in Tüfekci (2014) based on the prediction results. Finally, Section 5 concludes the paper.

2. CONTINUOUS CONDITIONAL RANDOM FIELD

C-CRF is one of the probabilistic graphical models used to express the complex dependence structure among system output variables by conditional probability distributions. A key advantage of the C-CRF is the flexibility to include a wide variety of arbitrary and non-independent features as inputs (McCallum, 2002). In addition, the C-CRF has a power of explanation because it is a probabilistic graphical model, as depicted in Figure 2, where each node denotes a variable and each edge de-

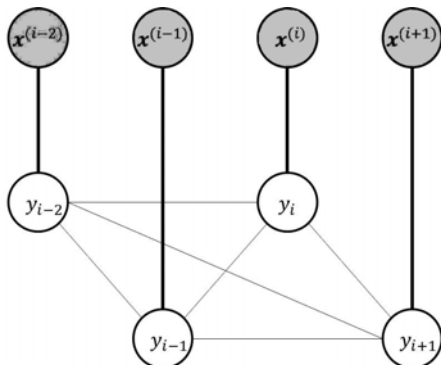


Figure 2. C-CRF structure.

notes a relationship between the two nodes. Here, $x^{(i)}$ and y_i are values of the i th input and output variables, respectively. We adopt the C-CRF structure depicted in Figure 2 to construct the proposed model.

A C-CRF can be represented as a conditional probability distribution $P(y|x)$ as follows:

$$P(y|x) = \frac{1}{z(x, \alpha, \beta)} \exp\left(\sum_{i=1}^n A(\alpha, y_i, x^{(i)}) + \sum_{i,j} I(\beta, y_i, y_j, x)\right) \quad (1)$$

where n is the number of observations, α and β are vectors of parameters, x is the vector of input variables, and $Z(x, \alpha, \beta)$ is a normalization factor that ensures $P(y|x)$ a proper probability distribution:

$$Z(x, \alpha, \beta) = \int \exp\left(\sum_{i=1}^n A(\alpha, y_i, x^{(i)}) + \sum_{i,j} I(\beta, y_i, y_j, x)\right) dy. \quad (2)$$

In (1), $A(\alpha, y_i, x^{(i)})$ and $I(\beta, y_i, y_j, x)$ are the association potential between the i th input variable vector $x^{(i)}$ and output variable y_i (represented as thick lines in Figure 2), and the interaction potential between outputs y_i and y_j (drawn in fine lines in Figure 2), respectively. The association potential indicates the relationship between inputs and outputs, while the interaction potential indicates the relationship among outputs. In the C-CRF application, the association and interaction potentials are often defined as linear combinations of fixed feature functions in terms of α and β as follows (McCallum, 2002):

$$A(\alpha, y_i, x^{(i)}) = \sum_{k=1}^{K_1} \alpha_k f_k(y_i, x^{(i)}), \quad (3)$$

$$I(\beta, y_i, y_j, x) = \sum_{k=1}^{K_2} \beta_k g_k(y_i, y_j, x). \quad (4)$$

Computing the normalization function in (2) is essential to obtain the exact probability noted in (1), but doing so is very complicated and sometimes intractable. In Radosavljevic *et al.* (2010), it is shown that if $A(\alpha, y_i, x^{(i)})$ and $I(\beta, y_i, y_j, x)$ are defined as quadratic functions in terms of y , then the sum $A(\alpha, y_i, x^{(i)}) + I(\beta, y_i, y_j, x)$ can be transformed into the form of $(y - \mu)^T \Sigma^{-1} (y - \mu) + constant$. This expression corresponds to a multivariate Gaussian distribution with a mean vector μ and a covariance matrix Σ . If the C-CRF is converted into the form of a multivariate Gaussian distribution, then the learning task of the parameters becomes relatively easier. A more detailed explanation can be found in Section 3.

Learning a C-CR Requires defining values of the parameters α and β , there by maximizing the conditional log-likelihood $L(\alpha, \beta) = \sum_{i=1}^n \log P(y_i | x^{(i)})$, and then

the result would be obtained as follows:

$$(\hat{\alpha}, \hat{\beta}) = \arg \max_{\alpha, \beta} (L(\alpha, \beta)). \quad (5)$$

Learning can be done by applying a standard optimization algorithm, such as the gradient ascent method. The inference task requires finding the output values for a given set of input values and estimated parameters, such that the conditional probability $\Pr(y|\mathbf{x})$ is maximized:

$$\hat{y} = \arg \max_y P(y|\mathbf{x}). \quad (6)$$

3. THE C-CRF MODEL FOR POWER GENERATION PREDICTION

In this section, we describe in detail the proposed C-CRF model for regression analysis in power generation. We introduce the same input and target variables as used in Tüfekci (2014). More specifically, ambient temperature (AT, measured in degrees in Celsius), atmospheric pressure (AP, in units of millibars), relative humidity (RH, measured as percentages), and exhausts steam pressure (or vacuum, V, in cmHg) are introduced as input variables in the dataset. As for the target variable, full load electrical power (PE, in megawatt) is used. All measurements are obtained through the sensors and are averaged hourly to provide the values of the input and target variables.

ANN, RT, and MLR were applied to predict the full load electrical power output in previous study (Al-Rashidi and El-Naggar, 2010). Employing the input variables as attributes, therefore, we introduce three feature functions to model the association potential that describes the dependency between the input and target variables for a given observation $\mathbf{x}^{(i)}$ as follows:

$$\begin{aligned} f_1(y_i, \mathbf{x}^{(i)}) &= -(y_i - \text{ANN}(\mathbf{x}^{(i)}))^2, \\ f_2(y_i, \mathbf{x}^{(i)}) &= -(y_i - \text{RT}(\mathbf{x}^{(i)}))^2, \\ f_3(y_i, \mathbf{x}^{(i)}) &= -(y_i - \text{MLR}(\mathbf{x}^{(i)}))^2, \end{aligned} \quad (7)$$

where $\text{ANN}(\mathbf{x}^{(i)})$, $\text{RT}(\mathbf{x}^{(i)})$, and $\text{MLR}(\mathbf{x}^{(i)})$ are the outputs of ANN, RT, and MLR, respectively. These feature functions agree with the basic principle for association potentials, that is, their values must increase for more accurate predictions. As a result, the following linear combination of these features provides insight on how much one can trust the prediction methods ANN, RT, and MLR based on the learned parameter vector $\alpha = (\alpha_1, \alpha_2, \alpha_3)$:

$$A(\alpha, y_i, \mathbf{x}^{(i)}) = -\alpha_1 (y_i - \text{ANN}(\mathbf{x}^{(i)}))^2 - \alpha_2 (y_i - \text{RT}(\mathbf{x}^{(i)}))^2 - \alpha_3 (y_i - \text{MLR}(\mathbf{x}^{(i)}))^2. \quad (8)$$

For example, a large α_2 places big penalty for mistakes in the RT model and therefore, each α_i acts as a quality indicator of the corresponding prediction method.

To model the interaction potential we introduce a feature function:

$$g(y_i, y_j, \mathbf{x}) = -S_{i,j} (y_i - y_j)^2, \quad (9)$$

where $S_{i,j}$ denotes a similarity between data i and j , and the corresponding interaction potential is given by:

$$I(\beta, y_i, y_j, \mathbf{x}) = -\beta S_{i,j} (y_i - y_j)^2. \quad (10)$$

The learned parameter β represents the level of correlation within neighboring outputs. That is, a large value of β implies a high correlation between y_i and y_j .

Finally, the resulting C-CRF model is given by:

$$P(y|\mathbf{x}) = \frac{1}{z(\mathbf{x})} \exp \left\{ \sum_i \sum_{k=1}^3 -\alpha_k (y_i - f_k(y_i, \mathbf{x}^{(i)}))^2 + \sum_{i,j} -\beta S_{i,j} (y_i - y_j)^2 \right\}. \quad (11)$$

We further map the distribution shown in Eq. (11) to a multivariate Gaussian distribution to reduce computational complexity. As one can see in Eq. (2), it is necessary to solve complicated integral calculus in order to obtain the C-CRF parameters. In Eq. (11), the potentials can be represented as quadratic forms as follows:

$$P(y|\mathbf{x}) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \times \exp \left(-\frac{1}{2} (y - \mu(\mathbf{x}))^T \Sigma^{-1} (y - \mu(\mathbf{x})) \right). \quad (12)$$

In the Gaussian mapping, the inverse of the covariance matrix Σ^{-1} is the sum of two $n \times n$ matrices, namely, $\Sigma^{-1} = 2(Q^1 + Q^2)$, where

$$\begin{aligned} Q_{i,j}^1 &= \begin{cases} \sum_{k=1}^3 \alpha_k, & \text{if } i = j, \\ 0, & \text{otherwise,} \end{cases} \text{ and} \\ Q_{i,j}^2 &= \begin{cases} \sum_j \beta S_{i,j} & \text{if } i = j, \\ -\beta S_{i,j}, & \text{otherwise.} \end{cases} \end{aligned} \quad (13)$$

Further, the mean vector $\mu(\mathbf{x})$ is computed as $\Sigma \theta$, where $\theta = 2 \sum_{k=1}^3 \alpha_k f_k(\mathbf{x})$.

With the multivariate Gaussian distribution that aims at maximizing the log-likelihood, the learning of the C-CRF in Eq. (5) becomes a convex optimization problem. As mentioned above, the gradient ascent method can be applied to learn the parameters. Specifically, we maximize the log-likelihood with respect to $\log \alpha_k$ and $\log \beta$ instead of α_k and β , which results in the new optimization problem becoming unconstrained. Derivatives of the log-likelihood function and updates of α and β in gradient ascent can be computed as follows:

$$\frac{\partial L}{\partial \log \alpha_k} = \alpha_k \frac{\partial L}{\partial \alpha_k}, \quad \frac{\partial L}{\partial \log \beta} = \beta \frac{\partial L}{\partial \beta}, \quad (14)$$

$$\log \alpha_k^{new} = \log \alpha_k^{old} + \eta \frac{\partial L}{\partial \log \alpha_k},$$

$$\log \beta^{new} = \log \beta^{old} + \eta \frac{\partial L}{\partial \log \beta}. \quad (15)$$

Now, the prediction shall be the expected value of Gaussian model, which is equal to the mean of the distribution:

$$\hat{y} = \arg \max_y P(y|\mathbf{x}) = \Sigma \boldsymbol{\theta}. \quad (16)$$

4. EXPERIMENT

In this section, an experiment to illustrate our model is provided. We describe the data used in the experiment, and then utilize the model to predict the probability distribution of the target value when the four input variables are given. Finally, the results are compared to existing methods to validate our model.

4.1 Data

We obtained a combined cycle power plant dataset from the UCI Machine Learning Repository (<http://archive.ics.uci.edu/ml/>). It is composed of 9,568 records collected while the combined cycle power plant was set to work with a full load over 674 days, contained four input variables (ambient temperature (AT), ambient pressure (AP), relative humidity (RH), and vacuum (V)), and a target variable (full load electrical power output (PE)). Table 1 shows the basic statistics of the dataset.

Table 1. Basic Statistics of Dataset

	AT (°C)	V(cmHg)	AP(mb)	RH(%)	PE(MW)
Min	1.81	25.36	992.89	25.56	420.26
Max	37.11	81.56	1033.30	100.16	495.76
Mean	19.65	54.31	1013.26	73.31	454.37
Standard deviation	7.45	12.71	5.94	14.60	17.07

Table 2. Correlations among Input Variables

	AT	V	AP	RH	PE
AT	1	-	-	-	-
V	0.84411	1	-	-	-
AP	-0.50755	-0.41350	1	-	-
RH	-0.54253	-0.31219	0.09957	1	-
PE	-0.94813	-0.86978	0.51843	0.38979	1

Table 2 contains the correlation matrix, indicating that the variables do not seem to be independent of each other.

4.2 Modeling and Evaluation

In order to validate and compare the performance of our model to those described in (Tüfekci, 2014), 5×2 cross-validation is applied. Root mean squared error (RMSE) and mean absolute error (MAE) are calculated and listed in Table 3 to assess the prediction accuracy, where:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad (17)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (18)$$

Euclidean similarity $S_{i,j}$ is employed to model the interaction potential, where $S_{i,j}$ between two records i and j is calculated as follows:

$$S_{i,j} = 1/(1 + \text{Euclidean Distance}(i, j)). \quad (19)$$

Note that when we calculate the Euclidean distance, in order to remove the scale effect of the variables, anormalized

Euclidean distance, $\sqrt{\sum_{k=1}^4 \left[\frac{1}{\text{std}(x_k)} (x_{ik} - x_{jk}) \right]^2}$, is used,

where $\text{std}(x_k)$ denotes the standard deviation of x_k , and x_{ik} denotes the value of variable k of the record i .

With the association and interaction potentials, we constructed a C-CRF model and transform it into a multivariate Gaussian distribution as in Eqs. (11) and (12). Parameters are estimated by applying Eqs. (14) and (15) with arbitrarily chosen initial values. The resulting estimated values of parameters are shown in Table 4.

Table 3. RMSE and MAE of ANN, RT, and MLR

	RMSE	MAE
ANN	4.074	3.205
RT	4.212	3.268
MLR	4.412	3.539

Table 4. Estimated parameters values

Parameter	Estimated Value
α_1	0.2002605
α_2	0.1500076
α_3	0.1496928
β	0.2408714

Table 5. Comparison results between our C-CRF model and those described in Tüfekci (2014)

	RMSE	MAE
Least Median Square	4.888	3.859
Support Vector Poly Kernel	4.887	3.859
K*	4.552	3.532
Bagging REP Tree	4.239	3.220
Model REP Tree	4.462	3.460
Model Tree Regression	4.428	3.428
REP Trees	4.518	3.424
Our C-CRF Model	3.978	2.970

We finally obtain the complete model in a multi-variate Gaussian distribution form as follows:

$$P(\mathbf{y}|\mathbf{x}) = \frac{1}{5513.672453} \exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu}(\mathbf{x}))^T \boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu}(\mathbf{x}))\right) \quad (20)$$

The RMSE and MAE of the predicted values obtained by Eq. (16) in our model are compared to those provided in the last two columns of Table 10 in (Tüfekci, 2014). The result is presented in Table 5:

Table 5 indicates that the C-CRF model presented in the paper shows the lowest RMSE and MAE (written in boldfaced numbers) among all models.

4.3 Application

The C-CRF model presented here is not only utilized to predict the value of power generation by CCP, but also used in various other ways. For example, it can be used to derive an interval estimation of the CCP's power generation and to calculate the probability distribution of it for a given set of input variables. We have selected three example problems to illustrate its applicability in this section.

As for the interval estimation and probability calculation, we randomly selected a record, i.e., record number 6663 in the dataset, whose feature values are: $AT_{6663} = 26.63^\circ\text{C}$, $AP_{6663} = 1,012.66\text{mb}$, $V_{6663} = 64.44\text{cmHg}$, $RH_{6663} = 61.19\%$, and $PE_{6663} = 442\text{MW}$. We also obtained $\mu(x_{6663}) = 459.45$, and $\sigma_{6663} = \sqrt{(\boldsymbol{\Sigma}^{-1})_{6663,6663}} = 21.46$. The result

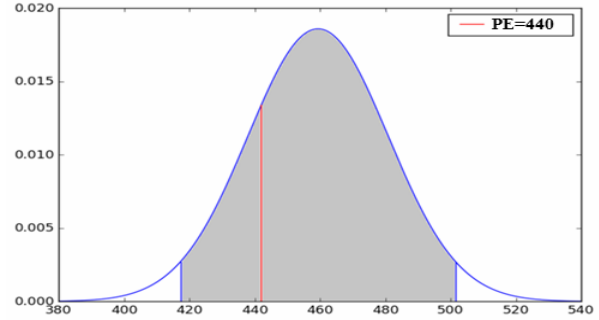


Figure 3. Probability density function of the PE of record 6663.

tant probability density function of PE (in MW) is as follows:

$$P(PE_{6663} | \mathbf{x}_{6663}) = \frac{1}{53.79} \exp\left(-\frac{(x - 459.45)^2}{921.0632}\right) \quad (21)$$

Since the mean and standard deviation are 459.45 and 21.46, respectively, the 95% confidence interval for the target value PE can be easily calculated as $459.45 \pm 1.96 \times 21.46 = (417.39, 501.51)$, which contains the real output value $PE_{6663} = 442\text{MW}$. A graph of the probability density function (blue line), corresponding to the confidence interval (hatched area) and the real output value (red line) are depicted in Figure 3.

The next example demonstrates the probability that the CCP's power generation is greater than 470MW when AT, AP, V, and RH are given as in the record 6663. It is formally described as:

$$\begin{aligned} & \Pr(PE_{6663} > 470\text{MW} | AT_{6663} = 26.63^\circ\text{C}, \\ & AP_{6663} = 1,012.66\text{cmHg}, V_{6663} = 64.44\text{cmHg}, \\ & RH_{6663} = 61.19\%). \end{aligned} \quad (22)$$

Since PE_{6663} follows a normal distribution with a mean of 459.45 and a standard deviation of 21.46, it is easily transformed into standard normal distribution form and calculated as:

$$\Pr\left(Z > \frac{470 - 459.45}{21.46} = 0.4916\right) = 0.6885, \quad (23)$$

where Z denotes the random variable which follows a standard normal distribution.

Lastly, we randomly selected two records, 143 and 867, from the dataset. Using our model, we calculated the probability that the PE of record 867 is greater than that of record 143, that is, $\Pr(PE_{143} < PE_{867})$, while utilizing the respective input variables for each record. The input variables of record 143 are: $AT_{143} = 23.02^\circ\text{C}$, $AP_{143} = 1,011.74\text{mb}$, $V_{143} = 59.21\text{cmHg}$, and $RH_{143} = 83.18\%$.

Input variables of record 843 are: $AT_{867} = 29.05^{\circ}\text{C}$, $AP_{867} = 1,011.33\text{mb}$, $V_{867} = 70.32\text{cmHg}$, and $RH_{867} = 72.50\%$. The probability density functions of PE_{143} and PE_{867} are presented in (24) and (25):

$$P(PE_{143} | \mathbf{x}_{143}) = \frac{1}{49.43} \exp\left(-\frac{(x - 432.66)^2}{777.7568}\right), \quad (24)$$

$$P(PE_{867} | \mathbf{x}_{867}) = \frac{1}{53.67} \exp\left(-\frac{(x - 456.71)^2}{916.7762}\right). \quad (25)$$

A simulation approach is subsequently adopted to solve the problem. 10,000 random numbers having the probability density functions presented in Eqs. (24) and (25) are generated, respectively, and two sets of random numbers are compared. More precisely, $D_1 = \{D_{1i} : i = 1, 2, \dots, 10,000\}$ and $D_2 = \{D_{2j} : j = 1, 2, \dots, 10,000\}$ are the sets of random numbers having the probability density functions presented in Eqs. (25) and (26), respectively. As a result, the probability $\Pr(PE_{143} < PE_{867})$ is calculated as follows:

$$\Pr(PE_{143} < PE_{867}) = \frac{\sum_{i=1}^{10,000} \sum_{j=1}^{10,000} I(D_{1i} > D_{2j})}{10,000 \times 10,000} \quad (26)$$

where $I(D_{1i} > D_{2j})$ is an indication function that returns 1 if $D_{1i} > D_{2j}$ and 0, otherwise. By means of this approach, $\Pr(PE_{143} < PE_{867})$ is calculated as 0.8551.

5. CONCLUSION

This study proposed a C-CRF model to predict the power of a CCPP at full load. Machine learning approaches were preferred to help ensure accurate prediction instead of thermodynamic approaches, which involve some assumptions with many nonlinear equations and require significant computational time and effort.

In this study, we have introduced three feature functions with machine learning methods to model the association potential, and a feature function with Euclidean similarity to model the interaction potential. With potentials represented by a combination of feature functions, the C-CRF model is transformed into a multivariate Gaussian distribution with which training and inference tasks were performed. The accuracies in terms of RMSE and MAE were compared to show that our model outperformed existing methods. The presented model can be applied to any regression application, such as interval estimation.

REFERENCES

AlRashidi, M. R. and El-Naggar, K. M. (2010), Long term

- electric load forecasting based on particle swarm optimization, *Applied Energy*, **87**(1), 320-326.
- Clifton, A., Kilcher, L., Lundquist, J. K., and Fleming, P. (2013), Using machine learning to predict wind turbine power output, *Environmental research letters*, **8**(2), 0204009.
- Fan, G. F., Peng, L. L., Hong, W. C., and Sun, F. (2016) Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression, *Neurocomputing*, **173**, 958-970.
- Kaya, H., Tüfekci, P., and Gürgen, F. S. (2012), Local and global learning methods for predicting power of a combined gas and steam turbine, in: *International conference on emerging trends in computer and electronics engineering*.
- Kesgin, U. and Heperkan, H. (2005), Simulation of thermodynamic systems using soft computing techniques, *International journal of energy research*, **29** (7), 581-611.
- Kumar, S. and Hebert, M. (2003), Discriminative random fields: A discriminative framework for contextual interaction in classification, in: *computer Vision, Proceedings Ninth IEEE International Conference on*, 1150-1157.
- Lafferty, J., McCallum, A., and Pereira, F. C. (2011), Conditional random fields: Probabilistic models for segmenting and labelling sequence data.
- Liu, Y., Carbonell, J., Klein-Seetharaman, J., and Gopalakrishnan, V. (2004), Comparison of probabilistic combination methods for protein secondary structure prediction, *Bioinformatics*, **20**(17), 3099-3107.
- McCallum, A. (2002), Efficiently inducing features of conditional random fields, in: *proceedings of the Nineteenth conference on Uncertainty in Artificial Intelligence*, 403-410.
- Niu, L. X. and Liu, X. J. (2008), Multivariable generalized predictive scheme for gas turbine control in combined cycle power plant, in: *Cybernetics and Intelligent Systems, IEEE Conference on*, 791-796.
- Prokop, L., Misak, S., Snasel, V., Platos, J., Krömer, P. (2013), Supervised learning of photovoltaic power plant output prediction models, *Neural Network World*, **23**(4), 321-338.
- Radosavljevic, V., Vucetic, S., and Obradovic, Z. (2010), Continuous Conditional Random Fields for Regression in Remote Sensing, in: *ECAI*, 809-814.
- Tüfekci, P. (2014), Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods, *International Journal of Electrical Power and Energy Systems*, **60**, 126-140.
- Xie, J. and Hong, T. (2015), GEF Com 2014 probabilistic electric load forecasting: An integrated solution with forecast combination and residual, simulation,

International Journal of Forecasting.

- Yadav, V. and Srinivasan, D. (2011), A SOM-based hybrid linear-neural model for short-term load forecasting, *Neurocomputing*, **74**(17), 2874-2885.
- Yu, F. and Xu, X. (2014), A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network, *Applied Energy*, **134**, 102-113.