Steam turbine power prediction based on encode-decoder framework guided by the condenser vacuum degree

Yanning Lu¹, Yanzheng Xiang², Bo Chen¹, Haiyang Zhu², Junfeng Yue¹, Yawei Jin¹, Pengfei He¹, Yibo Zhao², Yingjie Zhu², Jiasheng Si², Deyu Zhou²*

¹ Power Engineering Center, Jiangsu Frontier Electric Technology Co., Ltd., Nanjing, Jiangsu, China, ² School of Computer Science and Engineering, Key Laboratory of Computer Network and Information Integration, Ministry of Education, Southeast University, Nanjing, Jiangsu, China

* d.zhou@seu.edu.cn

Abstract

The steam turbine is one of the major pieces of equipment in thermal power plants. It is crucial to predict its output accurately. However, because of its complex coupling relationships with other equipment, it is still a challenging task. Previous methods mainly focus on the operation of the steam turbine individually while ignoring the coupling relationship with the condenser, which we believe is crucial for the prediction. Therefore, in this paper, to explore the coupling relationship between steam turbine and condenser, we propose a novel approach for steam turbine power prediction based on the encode-decoder framework guided by the condenser vacuum degree (CVD-EDF). In specific, the historical information within condenser operation conditions data is encoded using a long-short term memory network. Moreover, a connection module consisting of an attention mechanism and a convolutional neural network is incorporated to capture the local and global information in the encoder. The steam turbine power is predicted based on all the information. In this way, the coupling relationship between the condenser and the steam turbine is fully explored. Abundant experiments are conducted on real data from the power plant. The experimental results show that our proposed CVD-EDF achieves great improvements over several competitive methods. our method improves by 32.2% and 37.0% in terms of RMSE and MAE by comparing the LSTM at one-minute intervals.

Introduction

A stable electricity supply is an important guarantee for effective production. Thermal power generation is currently one of the most important power generation ways in the world. A complete thermal power generation system contains different equipment with different functions (e.g., circulating water pumps, condensers, steam turbines and other equipment, where the steam turbine is the power generation equipment and the condenser is the main auxiliary equipment). [1] has pointed out that accurate power forecasting is crucial for steam turbine control which is a complex and challenging task.
There exist many approaches for output power forecasting of turbines. Some approaches adopt machine learning methods, such as regression-based methods, to predict the output power of turbines. Furthermore, many neural network-based methods have been used to exploit the features of turbine operating data and predict the output power of turbines more accurately.

However, a notable drawback of the methods mentioned above is that they only focus on the output power prediction based on the turbine information individually while ignoring the correlation with the rest equipment of the power generation system. For example, they ignore the coupling relationship between the condenser (the front-end equipment of the steam turbine) and the steam turbine, which is contradictory to the practical scenario because some of the input factors are unavailable in reality (e.g., temporally condenser vacuum). In this paper, we explore the prediction of the steam turbine output power by introducing the coupling relationship with the condenser. A natural intermediary for coupling the condenser and the steam turbine is the condenser vacuum degree (i.e., an indicator reflecting the working status of the condenser and an important metric for the operation of the thermometric generating set.), which is a key factor for predicting the output power of the steam turbine. However, as an intermediate factor between the condenser and the steam turbine, the condenser vacuum degree varies dynamically based on the time and the condition status of the equipment [2–4], which brings difficulty in accurate modelling the vacuum degree temporally. In addition, since the different types of input variables for the condenser and the steam turbine, it is challenging to jointly model these two pieces of equipment by introducing the vacuum degree information into the output power prediction.

To overcome the above issues, in this paper, we propose an Condenser Vacuum Degree guided approach based on Encoder-Decoder Framework for steam turbine output power prediction (CVD-EDF). We model the condenser and steam turbine at the encoder and decoder, respectively. The encoder predicts the condenser vacuum degree dynamically and the decoder predicts the steam turbine output power at the target moment. In specific, we adopt the multi-layer LSTM as the basic architecture of the encoder and the decoder. Moreover, a connection module consisting of an attention mechanism and a convolutional neural network (CNN) is proposed to capture the local and global information within the encoder. All the information is further introduced into the decoder to predict the steam turbine output power at the target moment.

In summary, the main contributions of this paper are listed as follows:

- A novel condenser vacuum degree guided approach based on the Encoder-Decoder framework for steam turbine power prediction is proposed. To the best of our knowledge, our work is the first attempt to take into account the coupling relationship between the steam turbine and the condenser when predicting the steam turbine output power.

- Experimental results on the real data from the power plant show that our proposed CVD-EDF outperforms several competitive baselines. It achieves an improvement of 32.2% RMSE and an improvement of 37.0% MAE by comparing the LSTM at one-minute intervals.

**Related work**

**Encoder-Decoder framework**

The Encoder-Decoder framework is popular in artificial intelligence which consists of an encoder and a decoder. The encoder is a neural structure that extracts features from raw inputs...
and passes them to the decoder. The decoder is another neural structure that incorporates the features from the encoder and makes decisions for the task. In the beginning, the framework is widely used in the field of signal processing because of its ability to compress dimensions. [5] adopts the auto-encoder [6] for bio-signals compression and [7] propose to use convolutional auto-encoder for ECG signals compression. The auto-encoder is based on the encoder-decoder framework and it can learn a latent space in an unsupervised way. Later, the encoder-decoder framework became popular in the field of natural language processing [8–11], using it to process two tasks simultaneously at the encoder and decoder. [12] first applies it to neural machine translation. The framework is suitable for tasks that generate a sequential output and it can also be applied to other areas such as computer vision and speech processing [13–15]. However, there is no approach that adopts the Encoder-Decoder framework for output power prediction of the steam turbine. In this paper, we first apply the Encoder-Decoder framework to model the thermal power generation system. The reason is that we can process two tasks simultaneously in the encoder and decoder and introduce information from the encoder into the decoder in a flexible way. The condenser and the steam turbine are modelled in the encoder and decoder respectively.

**Output power prediction of turbines**

In general, the methods of predicting the output power of turbines can be divided into two categories according to their basic techniques and methodologies, i.e., machine learning approaches and deep learning approaches. For machine learning approaches, [16] adopts two non-parametric techniques based on the tilting method and monotonic spline regression to predict the power of the wind turbine. [17] proposes a non-linear regression model for wind turbine power curve approximation. Polynomial regression and exponential power curves have also been applied for output power prediction [18–20]. With the development of deep learning, many neural network-based methods are proposed for steam turbine power prediction. Artificial neural network (ANN) models representing the real power plant have been introduced into the steam turbine power prediction task [21, 22]. [23] adopts a long short-term memory network (LSTM) to forecast the wind turbine power and further uses the Gaussian mixture model (GMM) to analyze the error distribution characteristics of short-term wind turbine power forecasting. [24] proposes to adopt a neural network to establish accurate numerical simulators of the power plant units. However, all the approaches mentioned above have limitations because they only consider the information of the turbine and largely ignore the correlation with the rest equipment of the power generation system.

[25] takes into account not only the turbine but also the boiler when predicting the output power. They propose to utilize two ANN models, one for the boiler and one for the turbine, which are integrated to predict the power output from a coal-fired plant. However, our goal in this paper is to explore an approach from a different perspective, i.e., exploring the prediction of the steam turbine output power by introducing the coupling relationship with the condenser. [26] propose a novel hybrid framework for hotspot prediction which also contains LSTM and CNN, but this framework is markedly different from ours, which is based on an Encoder-Decoder framework. And we are the first to adopt the Encoder-Decoder framework for output power prediction of the steam turbine.

**Preliminary**

**LSTM**

Long short-term memory network (LSTM) [27] is a Recurrent Neural Networks (RNN) [28, 29] architecture which is widely used in output power prediction of the steam turbine. It has
significant advantages in processing time series data, which leverages time series dependencies between data. A standard RNN cannot bridge more than 5-10 time steps [30] and the reason is that the back-propagating error signal tends to grow or shrink with each time step [31].

LSTM is designed to handle long-time dependencies and the architecture of LSTM is shown in Fig 1. A common LSTM cell consists of a cell, an input gate, an output gate, and a forget gate. The three gates are composed of a sigmoid neural net layer and a pointwise multiplication operation. The outputs of gates are numbers range [0, 1]. They control the input, output and forgetting of past information of the cell respectively. These three gates regulate the flow of information into and out of the cell. Due to the structure of LSTM, it is able to access information at a more distant step.

At time step \( t \), the input of LSTM is \( x_t \) and the hidden states and the cell states at time step \( t - 1 \) are \( h_{t-1}, c_{t-1} \) respectively. The forget gate \( f_t \) decides what information should be discarded and it is calculated as follows:

\[
f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)
\]  

Where \( W_f \) is a trainable parameter and \( b_f \) is a bias vector. The input gate \( i_t \) determines what information should be stored in the cell states, and the new candidate value \( \tilde{c}_t \) is obtained:

\[
i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)
\]

\[
\tilde{c}_t = \text{tanh}(W_c[h_{t-1}, x_t + b_c])
\]

Where \( W_i, W_c \) are trainable parameters and \( b_i, b_c \) are bias vectors. The cell states \( c_t \) at time step \( t \) is calculated as follows:

\[
c_t = f_t \ast c_{t-1} + i_t \ast \tilde{c}_t
\]

And the output gate \( o_t \) determines what is going to be output, and the output \( h_t \) can be obtained:

\[
o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)
\]

\[
h_t = o_t \ast \text{tanh}(c_t)
\]

Where \( W_o \) is a trainable parameter and \( b_o \) is a bias vector. Following previous research, we adopt LSTM as the basic component for constructing our model in this paper.
In this paper, we explore the coupling relationship between the steam turbine and the condenser and propose a novel approach for steam turbine power prediction based on the encode-decoder framework guided by the condenser vacuum degree (CVD-EDF). We introduce the proposed CVD-EDF in detail in this section. The model architecture is shown in Fig 2, which consists of three parts:

1. Encoder: A LSTM is adopted to capture the historical information of the condenser operating conditions data, and the condenser vacuum degree of the target moment is predicted through a multi-layer perceptron network (MLP).

2. Connection module: An attention mechanism and a convolutional neural network (CNN) [32] are proposed to capture the local and global features of the hidden states from the encoder respectively at each step of the decoding process.

3. Decoder: The local, global features and the steam turbine operating condition data are concatenated as the input of the decoder. In this way, the information of the condenser vacuum is introduced into the decoder. Then, the history information of the steam turbine operating conditions data is captured by another LSTM. The initial hidden states and cell states of the decoder LSTM are initialized with the last hidden states and cell states of the encoder LSTM. The output power of the steam turbine at the target moment is predicted by fusing various information.

**Encoder**

In the encoder part, the condenser is modelled. The input of the encoder $X = [x_1, \ldots, x_t, \ldots, x_L]$ is the historical data sequence of the condenser operating conditions, where $t$ is the length of the data sequence.
of the sequence. The encoder extracts features from $X$ and predicts the vacuum degree at the target time step.

We apply a multi-layer LSTM as the encoder to extract time series patterns from the input. The initial hidden states $h_0$ and cell states $c_0$ of the encoder LSTM are initialized to zero. The hidden states $H_e = [h_1, \ldots, h_i, \ldots, h_t]$ and cell states $C_e = [c_1, \ldots, c_i, \ldots, c_t]$ of the last layer of the encoder can be obtained, mapping historical data from the original space to feature space:

$$H_e, C_e = LSTM(X, H_{e-1}, C_{e-1})$$  \hspace{1cm} (7)

The condenser vacuum degree $v_t$ at timestep $t$ can be calculated via:

$$v_t = W_2ReLU(W_1h_t)$$  \hspace{1cm} (8)

Where $W_1$ and $W_2$ are trainable parameters, and $h_t$ is the last encoder hidden state at timestep $t$. $ReLU$ denotes the ReLU [33] activate function.

**Connection module**

The Connection module is proposed to capture the vacuum degree information from the encoder and pass it to the decoder. An attention mechanism and a CNN are adopted to capture the local and global features from the encoder part. The extracted features will be used as input to the decoder.

For local feature $l_i$ at time step $i$, We can explicitly determine the semantic relevance between the decoder hidden state $h_{d-1}$ (Eq 15) at time step $i - 1$ and $H_e$ by calculating the dot product:

$$A_i = softmax(h_{d-1}^tH_e^T)$$  \hspace{1cm} (9)

The local feature $l_i$ at time step $i$ is calculated as follows:

$$l_i = A_iH_e$$  \hspace{1cm} (10)

For global feature $G$, we adopt a 1-D convolution to perform feature mapping on $H_e$:

$$h_{n+1}^i = f(w_iH_e^T_{n+1})$$  \hspace{1cm} (11)

$$o_i = max(h_{i+1}^{n+1})$$  \hspace{1cm} (12)

Where $s$ is the size of filters, $n$ is the stride of convolution, $w_i$ represents the parameters of the $i$-th filter and $f(.)$ denotes the activation function. Then, max pooling is applied to reduce the dimensionality of the convolution output $h_{i+1}^i$. We concatenate the output of all filters to obtain the global feature $G$:

$$G = [o_1, o_2, \ldots, o_m]$$  \hspace{1cm} (13)

Where $m$ is the number of filters.

**Decoder**

In the decoder part, the steam turbine is modelled. To take into account the coupling relationship between the condenser and the steam turbine, the local and global features extracted by the connection module are used as additional inputs to the decoder. The decoder fuses the multiple features and predicts the output power at the target time step. The historical data sequence $Y = [y_1, \ldots, y_i, \ldots, y_t]$ of the steam turbine operating conditions is concatenated with the local features $L = [l_1, \ldots, l_i, \ldots, l_t]$ and the global features $G$ and forming the input...
Another multi-layer LSTM is adopted as the decoder. The initial hidden states $h^d_0$ and cell states $c^d_0$ of the decoder are initialised with the last hidden states $h^e_t$ and cell states $c^e_t$ of the encoder LSTM. In this way, the condenser vacuum information can be introduced into the decoder. Then the hidden states $H^d = [h^d_1, ..., h^d_i, ..., h^d_t]$ and cell states $C^d = [c^d_1, ..., c^d_i, ..., c^d_t]$ of the last layer of the decoder can be obtained:

$$H^d, C^d = \text{LSTM}(I, H^d_{t-1}, C^d_{t-1})$$

The steam turbine power $p_t$ at timestep $t$ can be calculated via:

$$p_t = W_4 \text{ReLU}(W_3 h^d_t)$$

Where $W_3$ and $W_4$ are trainable parameters, and $h^d_t$ is the last hidden state at timestep $t$.

**Training loss**

We adopt the mean square error loss (MSE) to represent the error of the predicted condenser vacuum degree $v_t$, the predicted steam turbine output power $p_t$ with the real value:

$$L_1 = \frac{1}{N} \sum_{t=1}^{N} (v_t - V_t)^2$$

$$L_2 = \frac{1}{N} \sum_{t=1}^{N} (p_t - P_t)^2$$

where $P_t$ and $V_t$ represent the real condenser vacuum degree and the real steam turbine output power. $N$ is the number of training samples. The training objective is to minimize the total loss $\mathcal{L}$:

$$\mathcal{L} = L_1 + L_2$$

**Experiments**

In this section, we evaluate the effectiveness of CVD-EDF by comparing it to other approaches and ablating several design choices in CVD-EDF to understand their contributions.

**Dataset**

The data used for experiments are collected from a thermal power plant via sensors in Jiangsu Province of China, which range from July 1 to November 22, 2021, at a time interval of one minute.

For condenser operating conditions, the following real-time data needs to be collected: feedwater flow, circulating water inlet pressure, reheat steam temperature, reheat steam pressure, main steam temperature, main steam pressure, main steam flow rate, heat supply temperature, heat supply pressure, and heat supply flow rate. For steam turbine operating conditions, the following real-time data needs to be collected: supply line flow, supply line pressure, supply line temperature, reheat steam temperature, reheat steam pressure, main steam temperature
and main steam pressure. Moreover, real-time data on the vacuum degree of the condenser and the output power of the steam turbine also needs to be collected.

As shown in Table 1, The amount of data is 208,677, and we divided them into the training set and test set, containing 180000 and 28677 samples respectively.

### Experimental settings

The length of historical data sequence $t$ is set to 10. For CVD-EDF, the number of layers of the LSTM encoder and LSTM decoder is 2, and the hidden state dimension is set to 64. The whole model is trained by the Adam optimizer [34] with a learning rate of 1e-3. The number of epochs is 40 and the mini-batch size of the input is set to 32. The number of parameters in the model is 104,210 and the hyper-parameters are chosen based on the evaluation results from the test set.

### Metrics

Root mean squared error and mean absolute error are adopted to evaluate the overall performance.

1. Root Mean Squared Error (RMSE): It is a standard way to measure a model in predicting quantitative data. $y_t$ and $\hat{y}_t$ represent the ground truth and predicted value respectively, and RMSE is computed as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (y_t - \hat{y}_t)^2}$$  \hspace{1cm} (20)

2. Mean Absolute Error (MAE): It measures the average magnitude of the errors in a set of predictions:

$$MAE = \sum_{t=1}^{N} |y_t - \hat{y}_t|$$  \hspace{1cm} (21)

### Baselines

Some machine learning methods are chosen as the baselines, including several regression algorithms: Linear Regression, Ridge regression, LASSO regression, Elastic Net regression, Decision Tree regression and Xgboost model. And we also use LSTM as a baseline.

### Main results

Table 2 reports the overall performance of our model and baselines at one minute, one hour and one-day intervals. As shown in Table 2, it can be observed that:

1. Our proposed CVD-EDF significantly outperform other baseline models at all time intervals for both RMSE and MAE in steam turbine output power prediction. Compared with
LSTM, CVD-EDF achieves an improvement of 32.2% RMSE and an improvement of 37.0% MAE at one-minute intervals. It proves the effectiveness of our proposed CVD-EDF and the reason for the improvement is that CVD-EDF takes into the vacuum degree information when predicting the output power of the steam turbine. It also demonstrates that taking into account the coupling relationship with the condenser is helpful for the output power prediction of the steam turbine.

2. When the time interval is longer, the advantages of our proposed CVD-EDF are more obvious. It achieves an improvement of 63.7% RMSE and an improvement of 68.8% MAE at one-day intervals compared to LSTM.

3. Regarding baseline models, LSTM achieves the best performance both on RMSE and MAE at one-minute and one-hour intervals while Decision tree regression performs best at one-hour intervals. Among the regression algorithms, ridge regression and linear regression have similar performance and they perform better than other algorithms, but they still fall short of our proposed CVD-EDF.

Fig 3 is a timeline chart, it shows the prediction results of our proposed CVD-EDF for 300 consecutive moments samples in the test set at one-minute intervals. The blue line represents

### Table 2. Performance comparisons among several baselines at one minute, one hour and one-day intervals. The two metrics, RMSE and MAE, are in megawatts (MW). (”) represents “the smaller the better”.

| Model Name        | RMSE/MW(”) | MAE/MW(’) |
|-------------------|------------|-----------|
| **At one-minute intervals** |            |           |
| Linear Regression | 2.9658     | 2.3128    |
| Ridge regression  | 2.8914     | 2.2440    |
| LASSO regression  | 75.1732    | 60.4122   |
| Elastic Net regression | 75.1732    | 60.4122   |
| Decision Tree regression | 3.5969     | 2.5721    |
| Xgboost           | 3.3797     | 2.5487    |
| LSTM              | 2.5947     | 2.1656    |
| CVD-EDF           | 1.7597     | 1.3635    |
| **At one-hour intervals** |            |           |
| Linear Regression | 131.3898   | 102.3281  |
| Ridge regression  | 126.0740   | 98.4166   |
| LASSO regression  | 4406.8974  | 3550.1358 |
| Elastic Net regression | 4406.8974  | 3550.1358 |
| Decision Tree regression | 134.7658   | 93.8488   |
| Xgboost           | 162.2692   | 121.7480  |
| LSTM              | 124.2180   | 109.7862  |
| CVD-EDF           | 68.0847    | 53.6146   |
| **At one-day intervals** |            |           |
| Linear Regression | 2480.5151  | 2182.2465 |
| Ridge regression  | 2355.0108  | 2071.8719 |
| LASSO regression  | 86223.0892 | 63732.7707|
| Elastic Net regression | 86223.0892 | 63732.7707|
| Decision Tree regression | 1901.7868  | 1457.5911 |
| Xgboost           | 3068.4372  | 2495.9845 |
| LSTM              | 2666.4429  | 2539.8750 |
| CVD-EDF           | 968.0921   | 792.4782  |

https://doi.org/10.1371/journal.pone.0275998.t002
the actual steam turbine output power, while the red, yellow, green and purple lines represent the predicted steam turbine output power of our proposed CVD-EDF, ridge regression, LSTM and decision tree regression model respectively. It can be figured that:

1. our proposed CVD-EDF model is more consistent with the real steam turbine output power compared to decision tree regression, ridge regression and LSTM.

2. All the approaches have the ability to track the trend of the actual output power. However, the LSTM predictions are much lower than the actual steam turbine power almost all the time and the decision tree regression is not able to predict accurately during the time periods when the output power varies drastically. The ridge regression performs better than decision tree regression and LSTM, but the predictions are still inaccurate and more volatile. Our proposed CVD-EDF predictions are more accurate and informative.

Ablation study
To further evaluate the effectiveness of each component, we conduct some ablation experiments on our model at one-minute intervals. We performed three ablation experiments:

- The input of the decoder does not contain the global features captured by CNN (i.e., CVD-EDF w/o CNN).

- The input of the decoder does not contain the local features captured by the attention mechanism (i.e., CVD-EDF w/o Attention).

- The input of the decoder is the historical data sequence of the steam turbine operating conditions without the local features and global features. The initial hidden states and cell states of the decoder LSTM are initialized to zero. Therefore, no condenser vacuum information is introduced from the encoder to the decoder and the model degenerates to a simple LSTM model. (i.e., CVD-EDF w/o Attention and CNN).

The results are shown in Table 3, and they can be summarized in the table:
1. Compared with CVD-EDF w/o Attention and CNN, CVD-EDF w/o Attention achieves an improvement of 24.6% RMSE and an improvement of 34.1% MAE while CVD-EDF w/o CNN achieves an improvement of 29.7% RMSE and an improvement of 34.4% MAE. It demonstrates that both the local feature and the global feature are beneficial for the steam turbine output power prediction.

2. Compared with CVD-EDF, the RMSE and MAE for CVD-EDF w/o CNN, CVD-EDF w/o Attention drop a lot. It means that the CNN and the attention mechanism are able to capture different aspects of vacuum degree information, and the model performs best when the two are combined.

3. Compared with CVD-EDF w/o Attention, CVD-EDF w/o CNN achieves an improvement of 6.9% RMSE and an improvement of 0.5% MAE. It means that the local features are relatively more important than the global features. The reason is that, at each step of the decoding process, the attention mechanism is able to dynamically determine different parts of the hidden states of the encoder. In contrast to global features, it is able to filter some irrelevant information and keep the most important information as input to the decoder.

**Error analysis**

To further illustrate the performance of our approach, Fig 4 shows the prediction error of CVD-EDF for 300 test samples at one-minute intervals and Fig 5 shows the prediction error of

| Model Name                              | RMSE/MW() | MAE/MW() |
|-----------------------------------------|-----------|----------|
| CVD-EDF w/o Attention                   | 1.9573    | 1.4271   |
| CVD-EDF w/o CNN                         | 1.8231    | 1.4198   |
| CVD-EDF w/o Attention and CNN           | 2.5947    | 2.1656   |
| CVD-EDF                                 | 1.7597    | 1.3635   |

https://doi.org/10.1371/journal.pone.0275998.t003

---

**Fig 4. Steam turbine power prediction error comparison at one-minute intervals.**

https://doi.org/10.1371/journal.pone.0275998.g004
CVD-EDF for 300 test samples at one-hour intervals, compared with decision tree regression, ridge regression and LSTM model. It can be observed that:

1. CVD-EDF effectively tracks the true steam turbine power trend and the error of CVD-EDF is smaller than other baselines both at one-minute intervals and one-hour intervals. The reason is that our approach leverages the encoder-decoder framework to introduce vacuum degree information in the steam turbine output power prediction.

2. LSTM, decision tree regression and ridge regression can track the trend of the actual steam turbine power. However, LSTM outputs much lower predictions than the real steam turbine output power almost all the time, while the ridge regression outputs much higher predictions than the real steam turbine output power most of the time. The decision tree regression predictions have large errors at some time steps. Therefore, the predictions of LSTM, decision tree regression and ridge regression are unreliable in a practical scenario.

**Conclusion**

In this paper, we propose a novel approach for steam turbine power prediction based on the encode-decoder framework guided by the condenser vacuum degree (CVD-EDF), which for the first time explores the information on the coupling relationship between the steam turbine and the condenser. The condenser and steam turbine are modeled separately in the encoder part and the decoder part. In addition, a connection module which is composed of the attention mechanism and the CNN is proposed to capture the local and global information from the encoder. All of the information is introduced into the decoding process for accurate power prediction of the steam turbine. Experimental results on the real data collected from the power plant in Jiangsu Province of China show that the proposed method outperforms other competitive baselines. In the future, we will consider taking connected components of the thermal power generation system into account in the steam turbine power forecast, such as circulating water pumps.
Acknowledgments

The authors thank all those who have contributed to this work (Deyu Zhou, Jiasheng Si and Yanzheng Xiang conceived the idea and wrote the paper; Yanning Lu, Bo Chen, Junfeng Yue, Yawei Jin, Pengfei He provided data; Haiyang Zhu, Yibo Zhao, Yingjie Zhu conducted experiments.).

Author Contributions

Data curation: Bo Chen, Junfeng Yue, Yawei Jin, Yingjie Zhu.

Formal analysis: Haiyang Zhu, Yibo Zhao, Yingjie Zhu, Jiasheng Si.

Methodology: Yanning Lu, Yanzheng Xiang, Yawei Jin, Jiasheng Si.

Software: Yanning Lu, Yanzheng Xiang.

Supervision: Yanning Lu, Junfeng Yue, Pengfei He, Yibo Zhao, Jiasheng Si, Deyu Zhou.

Writing – original draft: Yanzheng Xiang.

Writing – review & editing: Yanning Lu, Jiasheng Si, Deyu Zhou.

References

1. Sun L, Liu T, Xie Y, Zhang D, Xia X. Real-time Power Prediction Approach for Turbine Using Deep Learning Techniques. Energy, 2021, 233:121130. https://doi.org/10.1016/j.energy.2021.121130

2. Xiaocheng Z. A Model for Predicting Vacuum in the Condenser based on Elman Neural Network by Using Particle Swarm Optimization Algorithm. Thermal Power Generation, 2010, 4:7–11.

3. Kumar H, Rahul, Verma S, Bera S. Analysis of Machine Learning algorithms for Prediction of Condenser Vacuum in Thermal Power Plant. International Conference on Electrical and Electronics Engineering, 2020, 778–783.

4. Lu K, Gao S, Xiangkun P, Meng X, Sun W, et al. Multi-layer Long Short-term Memory based Condenser Vacuum Degree Prediction Model on Power Plant. E3S Web of Conferences, 2019, 136:01012. https://doi.org/10.1051/e3sconf/201913601012

5. Sunil Kumar K, Shivashankar D, Keshavamurthy K. Bio-signals Compression Using Auto Encoder. Journal of Electrical and Computer Engineering, 2021, 2:424–433.

6. Rumelhart DE, Hinton GE, Williams RJ. Learning Representations by Back-propagating Errors. Nature, 1986, 323(6088):533–536. https://doi.org/10.1038/323533a0

7. Wang F, Ma Q, Liu W, Chang S, Wang H, He J, et al. A Novel ECG Signal Compression Method Using Spindle Convolutional Auto-encoder. Computer Methods and Programs in Biomedicine, 2019, 175:139–150. https://doi.org/10.1016/j.cmpb.2019.03.019 PMID: 31104703

8. Zhou D, Xiang Y, Zhang L, Ye C, Zhang QW, Cao Y. A Divide-And-Conquer Approach for Multi-label Multi-hop Relation Detection in Knowledge Base Question Answering. Conference on Empirical Methods in Natural Language Processing, 2021, 4798–4808.

9. Zhou L, Zhou Y, Corso JJ, Socher R, Xiong C. End-to-end Dense Video Captioning with Masked Transformer. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, 8739–8748.

10. Ding G, Chen M, Zhao S, Chen H, Han J, Liu Q. Neural Image Caption Generation with Weighted Training and Reference. Cognitive Computation, 2019, 11(6):763–777. https://doi.org/10.1007/s12559-018-9581-x

11. Zhang L, Zhou D, Lin C, He Y. A Multi-label Multi-hop Relation Detection Model based on Relation-aware Sequence Generation. Conference on Empirical Methods in Natural Language Processing, 2021, 4713–4719.

12. Sutskever I, Vinyals O, Le QV. Sequence to Sequence Learning with Neural Networks. Advances in Neural Information Processing Systems, 2014, 27.

13. Chen H, Ding G, Lin Z, Zhao S, Han J. Show, Observe and Tell: Attribute-driven Attention Model for Image Captioning. International Joint Conference on Artificial Intelligence, 2018, 606–612.

14. You Q, Jin H, Wang Z, Fang C, Luo J. Image Captioning with Semantic Attention. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, 4651–4659.
15. Wang X, Chen W, Wu J, Wang YF, Wang WY. Video Captioning via Hierarchical Reinforcement Learning. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, 4213–4222.

16. Mehrjoo M, Jozani MJ, Pawlak M. Wind Turbine Power Curve Modeling for Reliable Power Prediction Using Monotonic Regression. Renewable Energy, 2020, 147:214–222. https://doi.org/10.1016/j.renene.2019.08.060

17. Marčiukiaitis M, Žutautaitė I, Martišauskas L, Jokšas B, Gecevičius G, Sfetsos A. Non-linear Regression Model for Wind Turbine Power Curve. Renewable Energy, 2017, 113:732–741. https://doi.org/10.1016/j.renene.2017.06.039

18. Lydia M, Selvakumar AL, Kumar SS, Kumar GEP. Advanced Algorithms for Wind Turbine Power Curve Modeling. IEEE Transactions on Sustainable Energy, 2013, 4(3):827–835. https://doi.org/10.1109/TSTE.2013.2247641

19. Carrillo C, Montaño AO, Cidrás J, Díaz-Dorado E. Review of Power Curve Modelling for Wind Turbines. Renewable and Sustainable Energy Reviews, 2013, 21:572–581. https://doi.org/10.1016/j.rser.2013.01.012

20. Villanueva D, Feijoo A. Comparison of Logistic Functions for Modeling Wind Turbine Power Curves. Electric Power Systems Research, 2018, 155:281–288. https://doi.org/10.1016/j.apsr.2017.10.028

21. Mathioudakis K, Stamatias A, Tsalavoutas A, Aretakis N. Performance Analysis of Industrial Gas Turbines for Engine Condition Monitoring. Proceedings of the Institution of Mechanical Engineers Part A: Journal of Power and Energy, 2001, 215(2):173–184.

22. Mathioudakis K, Stamatias A, Bonataki E. Allocating the Causes of Performance Deterioration in Combined Cycle Gas Turbine Plants. Journal of Engineering for Gas Turbines and Power, 2002, 124(2):256–262. https://doi.org/10.1115/1.1426407

23. Zhang J, Yan J, Infield D, Liu Y, Lien Fs. Short-term Forecasting and Uncertainty Analysis of Wind Turbine Power based on Long Short-term Memory Network and Gaussian Mixture Model. Applied Energy, 2019, 241:229–244. https://doi.org/10.1016/j.apenergy.2019.03.044

24. Boccaletti C, Cerri G, Seyedar B. A Neural Network Simulator of a Gas Turbine with a Waste Heat Recovery Section. Journal of Engineering for Gas Turbines and Power, 2001, 123(2):371–376. https://doi.org/10.1115/1.1361062

25. Smrekar J, Pandit D, Fast M, Assadi M, De S. Prediction of Power Output of a Coal-fired Power Plant by Artificial Neural Network. Neural Computing and Applications, 2010, 19(5):725–740. https://doi.org/10.1007/s00521-009-0331-6

26. Khan SD, Alarabi L, Basalamah S. Toward Smart Lockdown: A Novel Approach for COVID-19 Hotspots Prediction Using a Deep Hybrid Neural Network. Computers, 2020, 9:99. https://doi.org/10.3390/computers9040099

27. Graves A. Long Short-term Memory. Supervised Sequence Labelling with Recurrent Neural Networks, 2012, 37–44. https://doi.org/10.1007/978-3-642-24797-2_4

28. Werbos PJ. Backpropagation through Time: What It Does and How to Do It. Proceedings of the IEEE, 1990, 78(10):1550–1560. https://doi.org/10.1109/5.58337

29. Williams RJ, Zipser D. A Learning Algorithm for Continually Running Fully Recurrent Neural Networks. Neural Computation, 1989, 1(2):270–280. https://doi.org/10.1162/neco.1989.1.2.270

30. Gers FA, Schmidhuber J, Cummins F. Learning to Forget: Continual Prediction with LSTM. Neural Computation, 2000, 12(10):2451–2471. https://doi.org/10.1162/089976600300015015 PMID:11032042

31. Hochreiter S, Bengio Y, Frasconi P, Schmidhuber J, et al. Gradient Flow in Recurrent Nets: The Difficulty of Learning Long-term Dependencies. IEEE, 2001, 237-243.

32. Fukushima K, Miyake S. Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Visual Pattern Recognition. Competition and Cooperation in Neural Nets, 1982, 267–285. https://doi.org/10.1007/978-3-642-46466-9_18

33. Nair V, Hinton GE. Rectified Linear Units Improve Restricted Boltzmann Machines. International Conference on Machine Learning, 2010, 807–814.

34. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. International Conference on Learning Representations, 2015.