Short-Term Multiple Load Forecasting Model of Regional Integrated Energy System Based on QWGRU-MTL

Songyao Wang and Zhisheng Zhang *

In order to improve the accuracy of the multiple load forecasting of a regional integrated energy system, a short-term multiple load forecasting model based on the quantum weighted GRU and multi-task learning framework is proposed in this paper. Firstly, correlation analysis is carried out using a maximum information coefficient to select the input of the model. Then, a multi-task learning architecture is constructed based on the quantum weighted GRU neural network, and the coupling information among multiple loads is learned through the sharing layer in order to improve the prediction accuracy of multiple loads. Finally, the PSO algorithm is used to optimize the parameters of the quantum weighted GRU. The simulation data of a regional integrated energy system in northern China are used to predict the power and cooling loads on summer weekdays and rest days, and the results show that, compared with the LSTM, GRU and single task learning QWGRU models, the proposed model is more effective in the multiple load forecasting of a regional integrated energy system.

Keywords: short-term multiple load forecasting; regional integrated energy system; quantum weighted GRU; neural network; multi-task learning

1. Introduction

In recent years, energy consumption has increased continuously and energy problems have become increasingly prominent [1]. The traditional single energy system can no longer meet the development of today’s economy and society, and the integrated energy system (IES) is in the ascendant. The regional integrated energy system (RIES) refers to the integrated energy system in a region, which integrates cooling, heating, electrical and gas energy. Its planning and dispatching considers the coupling of multiple energy sources, resulting in a higher energy utilization and a more flexible and reliable system operation [2]. The load forecasting of RIES is the basis of system operation and dispatching. Scientific and accurate forecasting is the premise of the safe and stable operation of RIES, so its necessity is self-evident. Obviously, the load forecasting of RIES cannot be limited to a single load, unlike the traditional power system. Cooling, heating, electrical and other multiple loads must be considered as a whole [3]. However, the multi-load in RIES has strong volatility and randomness, which puts forward higher requirements for load forecasting [4].

The neural network has been widely used in the field of load forecasting. The quantum neural network, as a new computing model combining quantum computation theory and the neural network model, has a better information processing ability and generalization ability and has been applied in the field of load forecasting. A load forecasting model of a double-layer quantum BP neural network is proposed in literature [5]. Combining the feature selection and quantum BP neural network, a short-term load forecasting model of the power system is constructed in the literature [6]. The above literature did not consider the combination of deep quantum neural networks and did not discuss the conversion formulas for the input of the model from a real state to a quantum state in depth [7].

Compared with traditional power load forecasting, the multiple load forecasting of RIES is still in its infancy. A stacked auto-encode extreme learning machine is used...
in the literature [2] for multivariate load predictions, but this model does not consider using recurrent neural networks to process time load series. Considering the influence of natural gas on the electrical load in the integrated energy system, the automatic encoder and decoder are adopted for feature extraction and load forecasting in the literature [6]; however, this literature only forecasts the electrical load, and multiple load forecasting is not involved. In literature [9], the Pearson correlation coefficient is used to analyze weather factors, and Kalman filter algorithm is used to predict power, cooling and heating loads, but the Pearson correlation coefficient cannot fully analyze nonlinear data. In literature [10], a chaos algorithm is adopted to improve PSO, which is used to optimize the parameters of the WNN model for multiple load forecasting. Furthermore, a multiple load forecasting model combining a wavelet packet decomposition and RNN is built in literature [11]. The cooling and heating loads are predicted by means of a kernel principal component analysis and GRNN neural network in literature [12]. The above neural network models adopt relatively uncomplicated structures of the hidden layer, and the deep neural networks are not considered to deal with the multivariate load coupling problem. A hard-sharing mechanism for multi-task learning parameters using a three-layer LSTM neural network is constructed in literature [13], and a multi-task learning sharing layer using a deep belief network is formed in literature [14]. The above applications of multi-task learning help the models to better cope with the multi-energy coupling problem, and have achieved good results.

Based on the existing research results, a multiple load forecasting model based on the quantum weighted gated recurrent unit neural network (QWGRU) and multi-task learning (MTL) is proposed in this paper. Firstly, the maximum information coefficient is used to analyze the correlation between multiple loads and the relevance between loads and weather factors to form the multiple input sequence. Secondly, the quantum weighted GRU neural network [15] is adopted as the sharing layer to construct a multi-task learning framework. Then, the optimal parameters of the quantum weighted GRU are selected by particle swarm optimization to realize multiple load forecasting. Finally, the effectiveness of the proposed model is verified by simulation analysis.

2. Correlation Analysis Method of Load and Influencing Factors

Multiple energy sources in RIES interact with each other through energy conversion equipment, and there are strong correlations between multiple loads [16]. At the same time, weather and other factors also have a greater impact on the loads. In order to adequately analyze the degree of interaction between multiple loads and the correlation between loads and weather factors, and to effectively select the input of the short-term multiple load forecasting model, a correlation analysis between multiple loads and weather factors is required. The Pearson method is mostly used in correlation analyses, but it is mainly used to analyze the correlation between linear variables, and performs poorly on nonlinear variables [17]. The maximum information coefficient (MIC) is a correlation analysis method based on mutual information. It is applicable to both linear and non-linear data, and the abnormal values in the data will not affect its analysis results [18]. Therefore, MIC is suitable for analyzing the non-linear relationship between the multi-load of RIES and influencing factors in this paper.

For a group of two-dimensional variables \((X, Y)\), \(m\) and \(n\) intervals are divided in their respective directions, and then the two-dimensional variables form a grid of \(m \times n\) on the plane. For the two-dimensional variables sample dataset \(D\), the MIC formulas are as follows [19]:

\[
MIC(X, Y) = \max_{m,n < B(N)} \left\{ M(D)_{m,n} \right\}
\]

\[
M(D)_{m,n} = \frac{I^*(D, m, n)}{\log_2 \min\{m, n\}}
\]

In the formulas, \(I^*(D, m, n)\) is the maximum mutual information of two-dimensional variables for data set \(D\), \(N\) is the number of samples and \(B(N)\) is a function of the sample
size, where its value is usually set to $N^{0.6}$ [20]. $MIC(X,Y)$ takes the value in $[0,1]$, and the closer the value of $MIC(X,Y)$ is to 1, the stronger the correlation between the two variables.

The data of the electrical load, cooling load and weather factors (temperature, solar radiation and humidity) in a RIES of a business office in the north of China in summer are used in this paper. Based on the above maximum information coefficient theory, the correlation between multiple loads and the relevance between loads and influencing factors in summer is analyzed. The results are shown in Table 1. $E$, $C$, $T$, $F$ and $HSL$ in the table represent the electrical load, cooling load, temperature, solar radiation and humidity, respectively.

Table 1. MIC value of each variable in summer.

| Features | $E$  | $C$  | $T$  | $F$  | $HSL$ |
|----------|------|------|------|------|-------|
| $E$      | 1    | 0.991| 0.454| 0.841| 0.204 |
| $C$      | 0.991| 1    | 0.51 | 0.565| 0.184 |

The results show that the MIC value between the electrical load and cooling load in summer is close to 1, which proves that there is a strong correlation between multiple loads of RIES that cannot be ignored in load forecasting. At the same time, the solar radiation, temperature and humidity also have a certain degree of influence on multiple loads, which should be taken into account in load forecasting.

3. QWGRU-MTL Load Forecasting Model

The quantum neural network has a stronger generalization capability and information processing ability, whereas multi-task learning can effectively learn the coupling information of the input. Combining the advantages of both, a QWGRU-MTL multiple load forecasting model is constructed in this paper.

3.1. GRU Neural Network Model

LSTM and GRU appear successively in order to deal with the long-term dependence problem that the simple recurrent neural network cannot solve. GRU is an improvement of LSTM and contains two gates: a reset gate and update gate. The update gate controls whether to retain the state of the previous moment and how much to retain. The reset gate can determine whether to combine the current state with the previous information. Its forward propagation formulas are as follows [21]:

$$r_t = \sigma(U_r h_{t-1} + W_r x_t)$$

$$z_t = \sigma(U_z h_{t-1} + W_z x_t)$$

$$\tilde{h}_t = \tanh(W x_t + U(r_t \odot h_{t-1}))$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_{t-1}$$

where: $r_t$ is the reset gate, $z_t$ is the update gate and $h_t$ is the output of the hidden layer. $W_r$ and $U_r$ are weights of the reset gate, $W_z$ and $U_z$ are weights of the update gate and $W$ and $U$ are weights of candidate hidden state. $\tanh$ is the hyperbolic tangent function, $\odot$ represents the Hadamard product of two matrices and $\sigma$ is the sigmoid function. The structure of GRU is shown in Figure 1.
is the activation function and \( \beta \) is the weight of \( x \). \( \alpha \) is the output of a neuron and \( F \) performs the internal product operation in this model.

\[
|\psi\rangle = [\cos \theta, \sin \theta]^T
\]

where \( x = [x_1, x_2, \cdots, x_m]^T \) is the input of a neuron, \( y \) is the output of a neuron and \( F \) performs the internal product operation in this model. \( |\phi_i\rangle = [\cos \alpha_i, \sin \alpha_i]^T \) is the weight quantum state and \( \alpha_i \) is its phase, and \( |\psi\rangle = [|\phi_1\rangle, |\phi_2\rangle, \cdots, |\phi_m\rangle]^T \) is the activity value quantum state and \( \beta \) is its phase. \( |\psi\rangle = [|\phi_1\rangle, |\phi_2\rangle, \cdots, |\phi_m\rangle]^T \) is the weighted quantum state vector. Thus, the input–output relationship of \( n \) quantum-weighted neurons is [15]:

\[
y = f(Wx)
\]

\[
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix} = f(Wx) \begin{bmatrix}
f(\sum_{i=1}^{m} x_i \langle \phi_1 | \phi_i \rangle) \\
f(\sum_{i=1}^{m} x_i \langle \phi_2 | \phi_i \rangle) \\
\vdots \\
f(\sum_{i=1}^{m} x_i \langle \phi_n | \phi_i \rangle)
\end{bmatrix} = f(Wx)
\]

where \( a \) and \( b \) are the probability amplitudes of a quantum state, are a pair of complex numbers and satisfy the normalization conditions, and \( |\rangle \) is the Dirac symbol, \( |0\rangle = [1, 0]^T \), \( |1\rangle = [0, 1]^T \).

Quantum weighted neurons simulate the transmission of transmitters between cells through four parts: weighting, aggregating, activating and stimulating. Figure 2 shows a diagram of the quantum weighted neuron model.

**Figure 1.** Structure diagram of GRU.

**Figure 2.** Diagram of quantum weighted neuron model.

### 3.2. Quantum Weighted Neuron

In the quantum computing system, a quantum state can be represented as follows [7]:

\[
|\psi\rangle = a|0\rangle + b|1\rangle = [a, b]^T
\]

where \( a \) and \( b \) are the probability amplitudes of a quantum state, are a pair of complex numbers and satisfy the normalization conditions, and \( |\rangle \) is the Dirac symbol, \( |0\rangle = [1, 0]^T \), \( |1\rangle = [0, 1]^T \).

Quantum weighted neurons simulate the transmission of transmitters between cells through four parts: weighting, aggregating, activating and stimulating. Figure 2 shows a diagram of the quantum weighted neuron model.
where \( |\phi(j)\rangle = [|\phi_{1j}\rangle, |\phi_{2j}\rangle, \ldots, |\phi_{mj}\rangle], |\phi_{ij}\rangle = [\cos \alpha_{ij}, \sin \alpha_{ij}]^T, |\varphi_i\rangle = [\cos \beta_j, \sin \beta_j]^T. \)

### 3.3. Quantum Weighted GRU Model

The above quantum weighted neurons are introduced into the input, reset gate, update gate and hidden layer states of the GRU model to form the QWGRU neural network model. Compared with the classical GRU, the introduction of quantum weighted neurons makes it have the superposition property of quantum states, and makes the network simulate the information processing mechanism of the biological nervous system more sufficiently. Each weight and activity value of the network will participate in the process of network optimization so that it can obtain a stronger information processing and optimization ability. At the same time, the real input and output form of the classical neural network is retained, which avoids the errors caused by the mutual conversion of the real state and quantum state. The forward propagation formulas of the QWGRU neural network with \( m - s - n \) structure are as follows:

\[
\begin{align*}
\begin{aligned}
    r_t^j &= \sigma(W_r x_t + U_r h_{t-1})^j = \sigma(F(x_t^T |\varphi_{wr}(j)|, |\varphi_{ur}(j)|) + F(h_{t-1}|\varphi_{ur}(j)|, |\varphi_{ur}(j)|)) \\
    &= \sigma(\sum_{i=1}^m x_i^j \cos((\alpha_{ur})_i j - (\beta_{ur})_i) + h_{t-1}^j \cos((\alpha_{ur})_i j - (\beta_{ur})_i))
\end{aligned}
\end{align*}
\]

\[
\begin{align*}
\begin{aligned}
    z_t^j &= \sigma(W_z x_t + U_z h_{t-1})^j = \sigma(F(x_t^T |\varphi_{zw}(j)|, |\varphi_{uw}(j)|) + F(h_{t-1}|\varphi_{uw}(j)|, |\varphi_{uw}(j)|)) \\
    &= \sigma(\sum_{i=1}^m x_i^j \cos((\alpha_{uw})_i j - (\beta_{uw})_i) + h_{t-1}^j \cos((\alpha_{uw})_i j - (\beta_{uw})_i))
\end{aligned}
\end{align*}
\]

\[
\begin{align*}
\begin{aligned}
    h_t^j &= \tanh(W_x x_t + U_r (z_t^j \otimes h_{t-1}))^j \\
    &= \tan(\sum_{i=1}^m x_i^j \cos((\alpha_{zw})_i j - (\beta_{zw})_i) + h_{t-1}^j \cos((\alpha_{uw})_i j - (\beta_{uw})_i))
\end{aligned}
\end{align*}
\]

where \( j = 1, 2, \ldots, s \). \( \varphi_{wr} \) is the weight quantum state constituting the weight matrix \( W_r \) and \( \varphi_{ur} \) is the active quantum state constituting the weight matrix \( W_r \). \( \alpha_{ur} \) is the phase of the \( \varphi_{ur} \) and \( \beta_{ur} \) is the phase of \( \varphi_{ur} \). Similarly, \( \varphi_{uw}, \varphi_{zw}, \varphi_{uw}, \varphi_{uw} \) are weight quantum states of weight matrices and \( \varphi_{uw}, \varphi_{zw}, \varphi_{uw}, \varphi_{uw} \) are quantum states of activity values, and \( \alpha_{uw}, \alpha_{zw}, \alpha_{uw}, \alpha_{uw}, \beta_{uw}, \beta_{uw}, \beta_{uw}, \beta_{uw} \) are their phase angles. The QWGRU model is shown in Figure 3.

![Diagram of QWGRU model.](image.png)

**Figure 3.** Diagram of QWGRU model.

### 3.4. Multi-Task Learning

Multi-task learning is an inductive transfer mechanism in the field of machine learning, with the main goal of improving the generalization ability of models [23]. The multi-task...
learning framework enables the simultaneous learning of multiple related tasks, each of which can share and complement the learning content through the sharing layer, thus enabling each task to learn input information better. In the feature sharing mechanism of multi-task learning, there are two ways: hard sharing and soft sharing. In the soft sharing mechanism, each task has its own model parameters, which can partially share the hidden layer, so the information sharing is more flexible. The structure of the hard sharing mechanism is relatively simple, and common features can be obtained by multiple tasks sharing the hidden layer, which is suitable for tasks with a strong coupling [24]. As a result of the strong coupling between multiple loads in RIES, the parameter hard sharing mechanism is adopted in this paper.

3.5. QWGRU-MTL Load Forecasting Model

In the traditional multi-task learning model, different tasks are separated only in the output layer, which makes it difficult to effectively learn their own unique characteristics. The multi-task learning model built in this paper deals with this problem. The QWGRU-MTL model built in this paper consists of an input layer, sharing layer, branch layer and output layer, and its structure is shown in Figure 4. The sharing layer is composed of several QWGRU neurons that are combined by recurrent structures. There are two branches in the branch layer, Tower A and Tower B, each of which is composed of several quantum weighted neurons described in Section 3.2. Based on Equation (12), the output of the tasks is derived as follows:

\[ T_A = \sigma(W_{TA}h_1)^k = \sigma(F(h_1^T|\varphi_{w_{TA}}(k))_r|\varphi_{w_{TA}})_k) = \sigma(\sum_{l=1}^{k} h_l^T \cos((\alpha_{w_{TA}})_l - (\beta_{w_{TA}})_k) \right) \]

\[ T_B = \sigma(W_{TB}h_1)^k = \sigma(F(h_1^T|\varphi_{w_{TB}}(k))_r|\varphi_{w_{TB}})_k) = \sigma(\sum_{l=1}^{k} h_l^T \cos((\alpha_{w_{TB}})_l - (\beta_{w_{TB}})_k) \right) \]

\[ y_t^A = \sigma(W_{y_A}T_A) = \sigma(F(T_A^T|\varphi_{w_{y_A}})_r|\varphi_{w_{y_A}}) = \sigma(\sum_{p=1}^{q} T_A^p \cos((\alpha_{w_{y_A}})_p - (\beta_{w_{y_A}}) \right) \]

\[ y_t^B = \sigma(W_{y_B}T_B) = \sigma(F(T_B^T|\varphi_{w_{y_B}})_r|\varphi_{w_{y_B}}) = \sigma(\sum_{p=1}^{q} T_B^p \cos((\alpha_{w_{y_B}})_p - (\beta_{w_{y_B}}) \right) \]

where \( k = 1, 2, \ldots, q \), \( T_A \) and \( T_B \) are the outputs of Tower A and Tower B, respectively and \( y_t^A \) and \( y_t^B \) are outputs of task A and task B, respectively. Refer to Sections 3.2 and 3.3 for the meaning of the parameters \( \varphi, \varphi, \alpha, \beta \).

![Figure 4. QWGRU-MTL structure diagram.](image-url)
In the training and forecasting process of the QWGRU-MTL model, firstly, the multi-task data are normalized to form the training set matrix and input it into the input layer of the model. Then, the data enter the sharing layer to learn common characteristics and coupling relations, followed by learning the unique characteristics of each task through the branch layer. Finally, the output of each task is obtained in the output layer. The mean square error function is used as the loss function, and each task firstly seeks its own loss, and then the weighted sum is carried out to obtain the total loss. The PSO algorithm [25] is adopted to optimize \( \alpha_{wr}, \alpha_{ur}, \alpha_{wz}, \alpha_{wy}, \alpha_{wTA}, \alpha_{wTB}, \alpha_{wTA}, \alpha_{wTB}, \beta_{wr}, \beta_{ur}, \beta_{wz}, \beta_{wy}, \beta_{wTA}, \beta_{wTB}, \beta_{wTA}, \beta_{wTB} \) in order to obtain the parameters with the smallest total loss, which are substituted into the model to realize short-term multiple load forecasting. The multiple load forecasting process is shown in Figure 5.

4. Example Simulation

The data set used in this paper includes the cooling load, electrical load and weather factors (temperature, solar radiation and humidity) for 365 days throughout a year, and the sampling step of each type of data is 1 h. The curves of the electrical load and cooling load in a certain period of summer are shown as Figure 6.
Energies 2021, 14, 6555 9 of 14

(a) Electrical load

(b) Cooling load

Figure 6. Load curves in summer. (a) Electrical load in a in a certain period of summer. (b) Cooling load in a in a certain period of summer.

4.1. Data Processing

Based on the results of the above correlation analysis, the input features are shown in Figure 7, where \( P_{d-1}^{k} \) represents the electrical load at \( k \) moment on the day before the forecast date, \( C_{d-1}^{k} \) represents the cooling load at \( k \) moment on the day before the forecast date and \( T_{d}^{k}, S_{d}^{k}, H_{d}^{k} \), respectively, represent the temperature, solar radiation and humidity at \( k \) moment on the day before the forecast date, and so on, for each of the symbols.

\[
\begin{array}{ccc}
\text{time} & k-1 & k & k+1 \\
\hline
d & T_{d-1}^{k} & T_{d}^{k} & T_{d+1}^{k} \\
& S_{d-1}^{k} & S_{d}^{k} & S_{d+1}^{k} \\
& H_{d-1}^{k} & H_{d}^{k} & H_{d+1}^{k} \\
d-1 & p_{d-1}^{k-1} & p_{d}^{k-1} & p_{d}^{k+1} \\
& c_{d-1}^{k-1} & c_{d}^{k-1} & c_{d}^{k+1} \\
d-2 & p_{d-2}^{k-1} & p_{d}^{k-2} & p_{d}^{k+1} \\
& c_{d-2}^{k-1} & c_{d}^{k-2} & c_{d}^{k+1} \\
d-3 & p_{d-3}^{k-1} & p_{d}^{k-3} & p_{d}^{k+1} \\
& c_{d-3}^{k-1} & c_{d}^{k-3} & c_{d}^{k+1} \\
\end{array}
\]

Figure 7. Features of the input.

4.2. Evaluating Indicator

Since this paper studies short-term multiple load forecasting, not only the single task, but also the overall effect of the forecasting model, needs to be analyzed; the mean absolute percentage error (\( \text{mape}_E \)), mean accuracy (\( \text{MA} \)) and weighted mean accuracy (\( \text{WMA} \)) are selected as evaluation indexes. Their calculation formulas are as follows:

\[
\text{mape}_E = \frac{1}{n} \sum_{k=1}^{n} |y_k - \hat{y}_k| \times 100\%
\]

\[
\text{MA} = \frac{1}{n} \sum_{k=1}^{n} \frac{|y_k - \hat{y}_k|}{|y_k|}
\]

\[
\text{WMA} = \alpha \text{MA} + (1-\alpha) \text{MA}_E
\]

where \( \alpha \) and \( 1-\alpha \) are the weights of the electrical load and cooling load, respectively. According to the data of the summer electrical load and cooling load, the critic weighting method in reference [26] is used to calculate them. The result shows that the weight of the electrical load is 0.54 and the weight of the cooling load is 0.46.

4.3. Results and Analysis

In order to verify the validity of the proposed QWGRU-MTL model, the QWGRU-MTL model, GRU model and LSTM model are used to forecast the electrical load and cooling load in summer. In the prediction of the weekday, the training set includes the cooling load, electrical load and weather factors of 19 weekdays, whereas, in the prediction of the rest day, the training set contains the cooling load, electrical load and weather factors of 17 rest days. The particle number of the PSO algorithm is set at 200, the range of speed is \([-0.36, 0.707]\), the range of linear decreasing inertia coefficient is \([0.4455, 1.172]\) and the learning rate \( c_1 = 1.85, c_2 = 2.15 \). The prediction effects of the three models are shown in Table 2.
4.2. Evaluating Indicator

Since this paper studies short-term multiple load forecasting, not only the single task, but also the overall effect of the forecasting model, needs to be analyzed; the mean absolute percentage error (\(E_{\text{mape}}\)), mean accuracy (\(MA\)) and weighted mean accuracy (\(WMA\)) are selected as evaluation indexes. Their calculation formulas are as follows:

\[
E_{\text{mape}} = \frac{1}{n} \sum_{k=1}^{n} \left( \frac{|Y_k - y_k|}{Y_k} \right) \times 100\% \tag{18}
\]

\[
MA = 1 - E_{\text{mape}} \tag{19}
\]

\[
WMA = \alpha_1 MA_1 + \alpha_2 MA_2 \tag{20}
\]

where \(\alpha_1\) and \(\alpha_2\) are the weights of the electrical load and cooling load, respectively.

According to the data of the summer electrical load and cooling load, the critic weighting method in reference [26] is used to calculate them. The result shows that the weight of the electrical load is 0.54 and the weight of the cooling load is 0.46.

4.3. Results and Analysis

In order to verify the validity of the proposed QWGRU-MTL model, the QWGRU-MTL model, GRU model and LSTM model are used to forecast the electrical load and cooling load in summer. In the prediction of the weekday, the training set includes the cooling load, electrical load and weather factors of 19 weekdays, whereas, in the prediction of the rest day, the training set contains the cooling load, electrical load and weather factors of 17 rest days. The particle number of the PSO algorithm is set at 200, the range of speed is \([-0.36, 0.707]\), the range of linear decreasing inertia coefficient is \([0.4455, 1.172]\) and the learning rate \(c_1 = 1.85, c_2 = 2.15\). The prediction effects of the three models are shown in Table 2.

| Weekday | Rest Day |
|---------|----------|
|         | Electrical Load | Cooling Load | Electrical Load | Cooling Load |
| \(E_{\text{mape}}/\%\) | QWGRU-MTL | 2.78 | 3.28 | 2.84 | 3.76 |
|         | GRU | 4.01 | 4.93 | 4.18 | 5.48 |
|         | LSTM | 3.45 | 4.67 | 3.91 | 4.78 |
| \(WMA/\%\) | QWGRU-MTL | 96.99 | | 96.74 | |
|         | GRU | 95.57 | | 95.22 | |
|         | LSTM | 95.99 | | 95.69 | |

It can be seen from the table that the weighted mean accuracy of the prediction results of the QWGRU-MTL model proposed in this paper on the weekday in summer can reach 96.99\%, which is 1.42\% and 1.00\% higher than that of the GRU model and LSTM model, respectively, and the weighted mean accuracy of the prediction results of the QWGRU-MTL model on the rest day in summer is 96.74\%, which is 1.52\% and 1.05\% higher than that of the GRU model and LSTM model, respectively. Focusing solely on the two sub tasks of electricity and cooling, the prediction error of the model proposed in this paper is reduced to a certain extent compared with the other two models. Compared with the GRU and LSTM models, the \(E_{\text{mape}}\) of the electrical load of the QWGRU-MTL model in the summer weekday is reduced by 1.23\% and 0.56\%, respectively, the \(E_{\text{mape}}\) of the cooling load is reduced by 1.65\% and 1.39\%, respectively, the \(E_{\text{mape}}\) of the electrical load in the summer rest day is reduced by 1.34\% and 1.07\%, respectively and the \(E_{\text{mape}}\) of the cooling load is reduced by 1.72\% and 1.02\%, respectively. At the same time, the accuracy of the cooling load of the three models is lower than that of the electrical load, because there is a large
gap in demand for cooling in different months, so the change law of the cooling load is not as easy to capture as the relatively stable electrical load.

Figures 8 and 9 show the forecasting curves of the electrical and cooling loads on a weekday and a rest day in summer. It can be seen from the diagram that the load forecasting curve of QWGRU-MTL proposed in this paper is closer to the actual load curve and that the forecasting effect is better. On the whole, the forecast errors of the three models are large near the peak of multiple loads. This is because there is more extreme weather in summer in the north of China than in other seasons, and the effect of meteorological factors on the peak loads is more obvious, which both makes the fluctuation of peak loads larger and makes it difficult to capture the change law, resulting in large forecast errors of peak loads.

In order to further analyze the advantages of multi-task learning, the prediction effects of the QWGRU-MTL model in this paper are compared with those of the quantum weighted GRU model using single-task learning (QWGRU-STL). The results of the comparison are shown in Table 3. Compared with single task learning, the multi-task learning model has a higher accuracy in forecasting electrical and cooling loads on the weekday and the rest day. The mean accuracy of the electrical load of QWGRU-MTL in the summer weekday is 1.68% higher than that of QWGRU-STL, and the mean accuracy of the cooling load of QWGRU-MTL is 1.93% higher than that of QWGRU-STL; the mean accuracy of the electrical load of QWGRU-MTL in the summer rest day is 2.72% higher than that of QWGRU-STL, and the mean accuracy of the cooling load of QWGRU-MTL is 2.62% higher than that of QWGRU-STL. This is because the QWGRU-STL model separates the relationship between the electrical load and cooling load. The input data only contain the information of a single load and lacks the coupling information with another load, so the model cannot fully learn effective information. However, the coupling characteristics between multiple loads can be used by QWGRU-MTL to improve the prediction accuracy.

---

### Table 2. Prediction effects of the three models.

| Models          | WMA/% Electrical Load | MA/% Electrical Load | WMA/% Cooling Load | MA/% Cooling Load |
|-----------------|------------------------|----------------------|-------------------|-----------------|
| QWGRU-MTL1      | 96.99                  | 2.72                 | 97.22             | 3.28            |
| QWGRU-MTL2      | 95.93                  | 3.45                 | 95.54             | 4.67            |
| LSTM            | 95.57                  | 3.91                 | 95.39             | 4.78            |
| QWGRU-STL       | 95.99                  | 2.78                 | 95.88             | 3.28            |

### Table 3. Comparison of multi-task learning.

| Models          | WMA/% Electrical Load | MA/% Electrical Load | WMA/% Cooling Load | MA/% Cooling Load |
|-----------------|------------------------|----------------------|-------------------|-----------------|
| QWGRU-MTL1      | 96.99                  | 2.72                 | 97.22             | 3.28            |
| QWGRU-MTL2      | 95.93                  | 3.45                 | 95.54             | 4.67            |
| LSTM            | 95.57                  | 3.91                 | 95.39             | 4.78            |
| QWGRU-STL       | 95.99                  | 2.78                 | 95.88             | 3.28            |

---

**Figure 8.** Forecasting curve of electrical load in summer.

**Figure 9.** Forecasting curve of cooling load in summer.
Table 3. Comparison of the effects of multi-task learning and single task learning.

| Models      | MA/|% | Weekday | Electrical Load | Cooling Load | Rest Day | Electrical Load | Cooling Load |
|-------------|----|---|---------|----------------|--------------|----------|----------------|--------------|
| QWGRU-MTL   | 97.22 | 96.72 | 97.16 | 96.24 |
| QWGRU-STL   | 95.54 | 94.79 | 94.44 | 93.62 |

In addition, a branch layer has been added into the multi-task learning framework in this paper. The model in this paper is recorded as QWGRU-MTL1, and the quantum weighted GRU multi-task learning model without a branch layer is recorded as QWGRU-MTL2. The prediction effects of the two models are shown in Table 4. As can be seen from Table 4, compared with the QWGRU-MTL2 model without the branch layer, the QWGRU-MTL1 model proposed in this paper improves the accuracy of short-term multiple load forecasting on the weekday in summer by 1.06%, and improves the accuracy of short-term multiple load forecasting on the rest day in summer by 1.73%, indicating that the branch layer plays a certain role in improving the prediction accuracy.

Table 4. Comparison of two multi-task learning models.

| Models      | WMA/|% | Weekday | Rest Day |
|-------------|-----|---|---------|----------|
| QWGRU-MTL1  | 96.99 | 96.74 |
| QWGRU-MTL2  | 95.93 | 95.01 |

Finally, the time required to train the above five models is shown in Table 5. Due to the fact that the particle swarm optimization algorithm needs repeated iterations in the parameter optimization, the model proposed in this paper needs a long training time, but it is still suitable for off-line learning and on-line prediction.

Table 5. The time required for training of five models.

| Models      | Training Time |
|-------------|---------------|
| QWGRU-MTL1  | 2489 s        |
| QWGRU-MTL2  | 1649 s        |
| QWGRU-STL   | 1706 s        |
| GRU         | 211 s         |
| LSTM        | 360 s         |

5. Conclusions

This paper presents a short-term multiple load forecasting model of RIES based on the QWGRU neural network and multi-task learning. After the simulation analysis on summer weekdays and rest days, the following conclusions are drawn:

(1) The multiple loads of RIES have a strong coupling, and the correlation between multiple loads and the relevance between loads and weather factors should be considered in the short-term multiple load forecasting;

(2) The quantum state superposition characteristics and multi-task learning mechanism of the QWGRU-MTL model can make the model fully learn the effective information in the input sequence and effectively improve the accuracy of multiple load forecasting. The simulation results show that, compared with other models, the proposed model has a higher forecast accuracy.

In the future, the structure of RIES will be more complex and the energy access will be more diverse. How to cope with the strong coupling and fluctuation of multiple loads of RIES in this case is a problem that needs further research in the future.
Author Contributions: Writing—original draft preparation, S.W.; writing—review and editing, Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 52077108.

Acknowledgments: This work was supported by National Natural Science Foundation of China (52077108).

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Li, K.; Sun, Y.; Li, S.; Ma, X.; Zhang, C. Load Forecasting Method for CCHP System Based on Deep Learning Strategy Using LSTM-RNN. In Proceedings of the 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA), Xi’an, China, 19–21 June 2019; pp. 827–831.
2. Wang, Y.; Ma, K.; Li, X.; Liang, Y.; Hu, Y.; Li, J.; Liu, H. Multi-Type Load Forecasting of IES Based on Load Correlation and Stacked Auto-Encode Extreme Learning Machine. In Proceedings of the 2020 10th International Conference on Power and Energy Systems (ICPES), Chengdu, China, 25–27 December 2020; pp. 585–589.
3. Luo, F.; Zhang, X.; Yang, X.; Yao, Z.; Zhu, L.; Qian, M. Load analysis and prediction of integrated energy distribution system based on deep learning. High Volt. Eng. 2021, 47, 23–32.
4. Liu, W.; Chen, X.; Sun, M.; Mei, C.; Li, Y.; Qi, M.; Gao, H. Multivariate Load Interval Prediction of Integrated Energy System Based on Multitask Learning. In Proceedings of the 2019 IEEE Sustainable Power and Energy Conference (iSPEC), Beijing, China, 21–23 November 2019; pp. 1084–1089.
5. Li, C. Study on Short Term Load Forecasting Based on Quantum Neural Network. Master’s Thesis, Hu Nan University, Changsha, China, 2011.
6. Li, P.; Yan, Y.; Zheng, W.; Jia, J.; Bai, B. Short-term load forecasting based on quantum neural network by evidential theory. Power Syst. Prot. Control 2010, 38, 49–53.
7. Zhu, Y. Research on Quantum Neural Network Model of Structure and Algorithm. Master’s Thesis, Northeast University, Shenyang, China, 2012.
8. Zhu, R.; Guo, W.; Gong, X. Short-term Power Load Forecasting Considering Correlations between Natural Gas and Power Load. Proc. CSU-EPSA 2019, 31, 27–32.
9. Zhao, F.; Sun, B.; Zhang, C. Cooling, Heating and Electrical Load Forecasting Method for CCHP System Based on Multivariate Phase Space Reconstruction and Kalman Filter. Proc. CSEE 2016, 36, 399–406.
10. Li, S.; Qi, J.; Bai, X.; Ge, L.; Li, T. A short-term load prediction of integrated energy system based on IPSO-WNN. Electr. Meas. Instrum. 2020, 57, 103–109.
11. Zhu, L.; Wang, X.; Ma, J.; Chen, Q.; Qi, X. Short-term load forecast of integrated energy system based on wavelet packet decomposition and recurrent neural network. Electr. Power Constr. 2020, 41, 131–138.
12. Ma, J.; Gong, W.; Zhang, Z. Short-term multiple load prediction model for regional integrated energy system based on Copula theory and KPCA-GRNN. Adv. Technol. Electr. Eng. Energy 2020, 39, 24–31.
13. Sun, Q.; Wang, X.; Zhang, Y.; Zhang, F.; Zhang, P.; Gao, W. Multiple load prediction of integrated energy system based on long short-term memory and multi-task learning. Autom. Electr. Power Syst. 2021, 45, 63–70.
14. Shi, J.; Tan, T.; Guo, J.; Liu, Y.; Zhang, J. Multi-task learning based on deep architecture for various types of load forecasting in regional energy system integration. Power Syst. Technol. 2018, 42, 698–707.
15. Xiang, W.; Li, F.; Wang, J.; Tang, B. Quantum Weighted Gated Recurrent Unit Neural Network and Its Application in Performance Degradation Trend Prediction of Rotating Machinery. Neurocomputing 2018, 313, 85–95. [CrossRef]
16. Chen, Z.; Gao, Z.; Chen, J.; Wu, X.; Fu, X.; Chen, X. Research on cooperative planning of an integrated energy system considering uncertainty. Power Syst. Prot. Control 2021, 49, 32–40.
17. Tian, H.; Han, A.; Yu, L.; Zhang, Z. Research on multi-load short-term forecasting model of regional integrated energy system based on GRA-LSTM neural network. Guangdong Electr. Power 2020, 33, 44–51.
18. Zhao, Y.; Wang, X.; Jiang, C.; Zhang, J.; Zhou, Z. A novel short-term electricity price forecasting method based on correlation analysis with the maximal information coefficient and modified multi-hierarchy gated LSTM. Proc. CSEE 2021, 41, 135–146.
19. Zhang, M.; Shen, X.; He, L.; Wang, H. Feature selection on maximum information coefficient for underwater target recognition. J. Northwestern Polytech. Univ. 2020, 38, 471–477. [CrossRef]
20. Reshef, D.N.; Reshef, Y.A.; Finucane, H.K.; Grossman, S.R.; McVean, G.; Turnbaugh, P.J.; Lander, E.S.; Mitzenmacher, M.; Sabeti, P.C. Detecting Novel Associations in Large Data Sets. Science 2011, 334, 1518–1524. [CrossRef] [PubMed]
21. Jung, S.; Moon, J.; Park, S.; Hwang, E. An Attention-Based Multilayer GRU Model for Multistep-Ahead Short-Term Load Forecasting. Sensors 2021, 21, 1639. [CrossRef] [PubMed]
22. Li, P. A learning algorithm and its applications to the quantum neural network model. Control Theory Appl. 2009, 26, 531–534.
23. Caruana, R. Multitask Learning. Mach. Learn. 1997, 28, 41–75. [CrossRef]
24. Xu, Z. Forum Duplicate Question Detection Based on the Wordto-Sentence Interaction Mechanism and Multi-Task Learning. Master’s Thesis, South China University of Technology, Guangzhou, China, 2020.
25. Wang, D.; Meng, L. Performance analysis and parameter selection of PSO algorithms. Acta Autom. Sin. 2016, 42, 1552–1561.
26. Song, D.; Liu, C.; Shen, C.; Shi, X.; Zang, L.; Feng, W. Multiple objective and attribute decision making based on the subjective and objective weighting. J. Shandong Univ. Eng. Sci. 2015, 45, 1–9.