Article

Daily Power Generation Forecasting Method for a Group of Small Hydropower Stations Considering the Spatial and Temporal Distribution of Precipitation—South China Case Study

Shaojun Yang, Hua Wei *, Le Zhang and Shengchao Qin

School of Electrical Engineering, Guangxi University, Nanning 530004, China; sjyang116@163.com (S.Y.); wjtclwwf@163.com (L.Z.); shengchao117@163.com (S.Q.)

* Correspondence: weihuagxu@163.com

Citation: Yang, S.; Wei, H.; Zhang, L.; Qin, S. Daily Power Generation Forecasting Method for a Group of Small Hydropower Stations Considering the Spatial and Temporal Distribution of Precipitation—South China Case Study. *Energies* **2021**, *14*, 4387. <https://doi.org/10.3390/en14154387>

Academic Editor: John M. Cimbala

Received: 24 April 2021

Accepted: 7 July 2021

Published: 21 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: This paper proposes a multimodal deep learning method for forecasting the daily power generation of small hydropower stations that considers the temporal and spatial distribution of precipitation, which compensates for the shortcomings of traditional forecasting methods that do not consider differences in the spatial distribution of precipitation. First, the actual precipitation values measured by ground weather stations and the spatial distribution of precipitation observed by meteorological satellite remote sensing are used to complete the missing precipitation data through linear interpolation, and the gridded precipitation data covering a group of small hydropower stations are constructed. Then, considering the time lag between changes in the daily power generation of the group of small hydropower stations and precipitation, the partial mutual information method is used to estimate the “time difference” between the two, and combined with the precipitation grid data, a data set of the temporal and spatial distribution of precipitation is generated. Finally, using only the temporal and spatial distribution of precipitation and historical power generation data, a multimodal deep learning network based on a convolutional neural network (CNN) and multilayer perceptron (MLP) is constructed, and a highly accurate prediction model for the daily power generation of small hydropower stations is obtained. Taking the real power generation data of a group of small hydropower stations in southern China as an example, after considering the temporal and spatial distribution of precipitation, the prediction accuracy of the proposed method is as high as 93%, which is approximately 5.8% higher than before considering the temporal and spatial distribution of precipitation. In addition, compared with mainstream methods such as support vector regression (SVR) and the long–short-term memory network (LSTM) (the average accuracy is about 87%), and the average accuracy improvement of the proposed method is approximately 6%.

Keywords: small hydropower stations; daily power generation forecasting; temporal and spatial distribution of precipitation; multimodal deep learning

1. Introduction

As a world-recognized renewable and clean energy, the rational development and utilization of small hydropower are of great significance for China to achieve the goal of “peak carbon and carbon neutrality”. According to the definition of relevant authorities in China, hydropower stations with an installed capacity of less than 50 MW are regarded as small ones [1]. By the end of 2019, China had built 45,445 small hydropower stations, with an installed capacity of 81,442 MW (equivalent to the total installed capacity of 3.5 Three Gorges hydropower stations), accounting for 22.9% of China’s installed hydropower capacity and 4.1% of China’s total installed power capacity. In 2019, China’s small hydropower generation capacity reached 253 million MWh, accounting for 19.5% of the total hydropower generation capacity and 3.5% of the total power generation capacity [2]. According to the

latest coal consumption standard of 308 g/kWh, the annual power generation of small hydropower stations in 2019 was equivalent to saving approximately 78 million tons of standard coal, reducing carbon dioxide emissions by approximately 195 million tons, and sulfur dioxide emissions by approximately 1.01 million tons. Small hydropower stations in China are mainly distributed in remote mountainous areas, and they are an important part of the local power supply. Small hydropower stations are characterized by decentralized development, local network formation, and local power supply, as well as advantages such as low construction cost, short construction period, and fast return; they are natural supplements to the main power grid and have irreplaceable advantages [3].

The large investment in small hydropower has made a great contribution to solving China's clean energy problems; however, most of the small hydropower stations are run-of-the-river stations, with a wide range of points and no adjustment capabilities. In practice, such stations are in a state of "power generation with water, but shutdown without water" in the original "blind" state. Their operation and output state are completely dominated by rainfall, with frequent fluctuations and uncertainties. With heavy rain, the power generation of small hydropower stations increases sharply, leading to a large and disorderly influx of small hydropower, causing a great impact on the main network, greatly destroying the power balance of the main network, and endangering the safety of the power system [4]. To realize optimal dispatching and resource allocation and ensure the safe, stable, and economic operation of the power system, it is necessary to make a reasonable small hydropower generation plan. Therefore, it is of great significance to accurately predict the generating capacity of small hydropower stations and provide a reference for the coordinated dispatch of multiple power sources in the power-dispatching department. Different from the well-developed forecasting models and methods of large and medium-sized hydropower stations, small hydropower generation capacity prediction research started late and faced a shortage of data (especially meteorological data and runoff data most relevant to small hydropower generation), large numbers, and poor model versatility. It is often difficult to learn from the mature hydrology and power generation forecasting methods of large- and medium-sized hydropower stations.

In view of the above problems, a few scholars have carried out relevant research. Paper [5] took the entire area of small hydropower as the object and assimilated snow observations to improve the forecasting results of power generation capacity of small hydropower in snow-rich areas. Paper [6] decomposed the power generation load of small hydropower into meteorological load components and long-term trend load components and predicted the power generation load of small hydropower by establishing a regression relationship between meteorological factors and meteorological load and a prediction model of long-term trend load characteristics. Paper [7] introduced the concept of regional synchronization characteristics of small hydropower generation for the modeling of small hydropower output based on the differences in climatic conditions in different regions. Paper [8] analyzed the influencing factors and characteristics of small hydropower generation and found that small hydropower output has a strong correlation with long-term precipitation, but the correlation with temperature and wind is very weak. When there is plenty of rainfall in summer, the output of small hydropower fluctuates greatly, and the change has a lag effect. Paper [9] inputted different combinations of rainfall and power generation for t days before the forecast day into the echo state network (ESN), obtained the prediction results of each combination, and used the comparison method to obtain the optimal result of the model. Papers [10,11] found that there are similar hydrological and meteorological conditions between large and small hydropower stations in the same basin. Therefore, the power generation of small hydropower stations can be predicted by using the predicted inflow of large- and medium-sized hydropower stations.

In summary, the key to improving the forecast accuracy of the power generation capacity of small hydropower stations lies in obtaining more abundant precipitation and historical operating data and adopting appropriate methods to fully extract the potential characteristic information. Considering the scattered distribution of small hydropower

stations, the wide area of rain collection, and the extremely uneven spatial distribution of precipitation, it is necessary to consider the difference in the temporal and spatial distribution of precipitation in forecasting the short-term power generation of small hydropower stations. However, no scholars have conducted relevant research on the abovementioned problems. Therefore, with the help of the precipitation distribution field obtained by meteorological satellite remote sensing observations and the partial mutual information method, the difference in the temporal and spatial distribution of precipitation is included in a forecast of the short-term power generation of small hydropower stations for the first time in this paper. First, the spatial distribution of precipitation observed by satellite remote sensing and the actual precipitation observed at ground meteorological stations are used to generate the precipitation grid covering the region of the group of small hydropower stations. Then, the partial mutual information method is used to select the precipitation time scale that has the most significant impact on changes in short-term power generation. Finally, combined with the recent major trends in historical power generation data, a model is built to forecast the short-term power generation of the group of small hydropower stations.

To fully explore the characteristic information contained in the temporal and spatial distribution of precipitation data and the general trend of recent and historical power generation, this paper proposes a multimodal deep learning network based on a convolutional neural network and multilayer perceptron (CM-MDLN). The convolutional neural network (CNN) is good at processing data in the form of multiple arrays and can deeply express the effective spatial feature information contained in the grid data of precipitation. Multilayer perceptron (MLP) has a simple structure and strong adaptive ability. It can theoretically approximate any non-linear function and can learn representative trend characteristics of recent changes from historical power generation data. In this paper, the effectiveness of the proposed method is verified by using real data from a group of small hydropower stations in southern China for approximately 3 years, and the results are compared with the results of other forecasting models.

2. Methodology

2.1. Precipitation Distribution Estimation Based on Satellite Remote Sensing

Precipitation is the most important factor affecting the power generation capacity of small hydropower stations. The accuracy of precipitation data directly affects the accuracy of daily power generation forecasts for small hydropower stations. Rain gauges are a traditional way of directly observing precipitation on the ground, but it is difficult to accurately capture the temporal and spatial distribution of precipitation due to the influence of the density of the station network and its spatial distribution. The spatial variations in precipitation are large, and small hydropower stations are scattered and remotely located. Therefore, it is difficult to predict the daily power generation of small hydropower stations with high precision by only using the actual precipitation values observed by rain gauges [12].

With the development and advancement of satellite remote sensing technology, quantitative precipitation observation (QPE) data based on satellite remote sensing inversion have become a new source of precipitation data. Satellite QPE data usually have broad coverage and high temporal and spatial resolution, which effectively compensates for the shortcomings of the spatial distribution of traditional meteorological station observations and can provide a new data reference for regions lacking meteorological data [13,14].

Satellite observations of precipitation are affected by radar sensor designs, weather conditions, and inversion algorithms, and they are not sufficient to accurately and quantitatively represent the precipitation values in the monitored region, but they can approximately reflect the spatial distribution of precipitation. Therefore, it is possible to combine the spatial distribution field of precipitation observed by satellites and the actual values of ground precipitation observed by rain gauges and use the method of linear interpolation

to estimate the approximate precipitation values are the locations without rain gauges, as shown in Figure 1.

Figure 1. Schematic diagram of precipitation estimation at any location.

First, we define the reference ratio of precipitation for a certain day as:

$$P_B = \frac{P_R}{P_S}, \quad (1)$$

where P_R represents the precipitation value observed by the ground meteorological station and P_S represents the precipitation value observed by the satellite at the longitude and latitude of the aforementioned ground meteorological station. Then, the actual precipitation at any location can be estimated as:

$$P_R(long, lat) = P_S(long, lat) \times P_B, \quad (2)$$

where *long* and *lat* represent the longitude and latitude, respectively, corresponding to the estimated position.

2.2. Hysteresis Effect of Precipitation

The area where the group of small hydropower stations is located has a wide rain-collecting surface and complex terrain, so it takes a long time for the precipitation in a certain period to have a significant impact on the output of the group of small hydropower stations. Therefore, the variation in the daily power generation of the small hydropower stations has an obvious time lag relative to the precipitation in a certain period, and its variation may be affected by the precipitation in the current period, the previous period, or even the previous several periods. When selecting the precipitation to forecast the daily power generation of small hydropower stations, it is necessary to consider not only the precipitation in the current period but also the precipitation patterns in previous periods.

Partial mutual information (PMI) is an extension of mutual information (MI) [15] and is used to measure the correlation between multiple random variables. PMI is similar to the partial correlation coefficient, which quantifies the dependency between the output Y and the variable Z after considering the input variable X [16–18].

When considering the influence of precipitation patterns in previous periods on the daily power generation of small hydropower stations, it is assumed that X and Z are the precipitation in two periods, and Y is the change in power generation after being affected. If there is a connection between X and Z , the correlation between Y and Z will be overestimated. Therefore, we use the conditional expectation to remove the information of

X contained in Y and Z before measuring their correlation. After removing the information of X , Y and Z are, respectively, denoted as u and v :

$$E[x|z] = \frac{1}{n} \sum_{i=1}^n w_i \left(x_i + (z - z_i)^T S_{zz}^{-1} S_{xz} \right) \quad (3)$$

$$w_i = \exp \left(-\frac{(z - z_i)^T S_{zz}^{-1} (z - z_i)}{2h^2} \right) / \sum_{j=1}^n \exp \left(-\frac{(z - z_j)^T S_{zz}^{-1} (z - z_j)}{2h^2} \right), \quad (4)$$

where S_{xz} is the cross-covariance between two random vectors X and Z , S_{zz} is the sample covariance of Z , and n is the number of samples in the vector.

$$u = Y - E[Y|Z] \quad (5)$$

$$v = X - E[X|Z] \quad (6)$$

Then, the partial mutual information between Y and Z , that is, the correlation between them, is denoted as:

$$PMI(Y, Z) = I(v, u) \approx \frac{1}{n} \sum_{i=1}^n \log \left[\frac{f(v_i, u_i)}{f(v_i)f(u_i)} \right], \quad (7)$$

where f represents the density estimation function, and we use the Gaussian kernel density estimation function to estimate the probability density of each variable.

The precipitation data set of the previous m periods is $C = \{P_t, P_{t-1}, P_{t-2}, P_{t-3}, \dots, P_{t-m}\}$, the change in power generation is $Y = \{\Delta E_t\}$, and the precipitation in the period with the most significant correlation with Y is stored in set S . Then, the PMI variable selection process is shown in Algorithm 1. The Akaike information criterion (AIC) [19] is used as the stopping criterion in the selection process, and its expression is as follows:

$$AIC = n \log \left(\frac{1}{n} \sum_{i=1}^n u_i^2 \right) + 2(p+1)^2, \quad (8)$$

where u is the residual of Y obtained after calculating the conditional expectation according to the selected variables, n is the number of samples, and p is the number of the selected precipitation periods. With the screening of independent variables, the AIC value decreases continuously, and the screening ends when it reaches the minimum value, which means that the input variable set with the most significant correlation has been selected.

2.3. Multimodal Deep Learning

The problem of forecasting the daily power generation of small hydropower stations is essentially a supervised regression problem. The mathematical explanation is to find a mapping function that satisfies $f(X) = Y$, where X is the feature vector composed of all factors affecting the power generation in the period to be forecasted, and Y is the power generation on the day to be forecasted.

This paper uses the spatial distribution of precipitation and historical power generation as predictive variables, which come from different information sources, represent different physical properties, have different data dimensions, and have obvious heterogeneous characteristics, which are called multimodal information in machine learning problems [20–22].

In this paper, we build a late fusion multimodal deep learning network based on a CNN and an MLP, which aims to jointly represent data of different modalities, capture the internal associations between different modalities, achieve information complementation through the mutual fusion of multimodal information, and improve the accuracy of the forecasting result.

Algorithm 1: The selection process of precipitation lag period.

Require: Input variable set C , output variable Y , optimal input variable set S , the number of input variables m

```

1: Initialize  $S \leftarrow null$ 
2: for  $i \leftarrow 1$  to  $m$  by 1 do
3:   Count the number of variables in the existing  $C$  in the  $i$ -th
     iteration, denoted by  $num(i)$ 
4:   for  $j \leftarrow 1$  to  $num(i)$  by 1 do
5:     if  $S = null$ 
6:        $u \leftarrow Y, v \leftarrow C(j)$ 
7:     else
8:        $u \leftarrow Y - E(Y|S), v \leftarrow C(j) - E(C(j)|S)$ 
9:     end
10:    Calculate the mutual information between  $u$  and  $v$ , de-
      noted by  $I(j)$ 
11:  end
12:  Calculate the  $AIC(i)$  of  $u$  in the  $i$ -th iteration
13:  Record the index of the maximum value in  $I$ , denoted by
      $index$ 
14:  if  $i \neq 1$ 
15:    if  $AIC(i) > AIC(i-1)$ 
16:      break
17:    end
18:  end
19:   $S(i) \leftarrow C(index)$ 
20:   $C \leftarrow \text{Delete } C(index) \text{ from } C$ 
21: end

```

2.4. Power Generation Forecasting Architecture

The overall architecture of the forecasting method for the daily power generation of the group of small hydropower stations proposed in this paper is divided into three parts: a multimodal data set, a multimodal deep learning network, and a late fusion network, as shown in Figure 2.

- (1) Multimodal data set. It is a heterogeneous mixed-dimensional data set composed of precipitation data with two-dimensional spatial distribution characteristics and daily power generation data with one-dimensional temporal series characteristics.
- (2) Multimodal network. The multimodal network is composed of a CNN and an MLP. The branch of the CNN has six layers ($L_{1,1}$ to $L_{1,6}$), and the input layer $L_{1,1}$ receives the data set of the spatial distribution of precipitation. The deep convolution layer $L_{1,2}$ to $L_{1,4}$ extracts the deep features of the input variables, and the ReLU function and the BatchNormalization method are used in each layer to activate and adjust the distribution of the extracted feature data. The flattening layer $L_{1,5}$ and the fully connected layer $L_{1,6}$ integrate the highly abstracted feature data after multiple convolutions to facilitate subsequent feature fusion. The MLP branch has three layers ($L_{2,1}$ to $L_{2,3}$) which receive the historical power generation data of the several previous periods to extract the variation characteristics of the power generation capacity of the entire group of small hydropower stations in the short term during the present period.
- (3) Late fusion. The late fusion network consists of four layers (L_1 to L_4), and the number of nodes in each layer is 16, 8, 4, and 1. The joint layer L_1 receives and dimensionally connects the temporal and spatial feature information of precipitation and the feature information of power generation capacity changes extracted from the multimodal network. The fully connected layers L_2 to L_4 form a simple neural

network, which carries out regression fusion and analysis on the extracted feature information and utilizes a linear function to activate the output layer to obtain the final forecasting value.

The detailed parameters of each layer of network in the proposed model are shown in Table 1.

Table 1. The parameters of each layer of the method presented in this paper.

	Layers											
	$L_{1,2}$	$L_{1,3}$	$L_{1,4}$	$L_{1,5}$	$L_{1,6}$	$L_{2,1}$	$L_{2,2}$	$L_{2,3}$	L_1	L_2	L_3	L_4
Num. of neurons	32	64	128	16	8	32	16	8	16	8	4	1
Size of conv. kernel	(2, 2)	(2, 2)	(2, 2)	/	/	/	/	/	/	/	/	/
Pooling block size	(1, 2)	(1, 2)	(1, 2)	/	/	/	/	/	/	/	/	/
Activation function	relu	relu	relu	relu	relu	relu	relu	relu	/	relu	relu	liner

Figure 2. The overall architecture of the forecasting method for the daily power generation of small hydropower stations.

3. Data Preprocessing

3.1. Data Description

We use the daily power generation data and precipitation data of small hydropower stations in Hechi city (HC) and Guilin city (GL) in southern China for approximately 3 years to verify the effectiveness of the proposed CM-MDLN method based on a CNN and an MLP. HC and GL have a large rain collection area, many large and small rivers, and the average annual rainfall is between 1200 and 1600 millimeters. It is a typical region with abundant hydropower resources.

3.2. Grid Division of the Spatial Distribution of Precipitation

The QPE data used in this paper are from the Global Precipitation Measurement (GPM) program satellite constellation. The GPM satellite was launched by NASA and JAXA in 2014, aiming to measure rainfall and snowfall on Earth with advanced radar and radiometers carried by the satellite [23]. The GPM core platform carries the first satellite-borne Ku/Ka-band dual-frequency precipitation radar (DPR) and a multi-channel GPM microwave imager (GMI). The minimum threshold value of precipitation measurement is 0.5 mm/h, which can achieve a good measurement of precipitation.

In HC City and GL City, there are 8 and 12 ground meteorological stations, respectively, and their location distribution is shown by the orange dots on the left side of Figures 3 and 4, respectively. Compared with the rain-collecting surface of tens of thou-

sands of square kilometres, the number of meteorological stations is particularly scarce. In accordance with the specifications of the $0.3^\circ \times 0.3^\circ$ (latitude and longitude) grid, we generate a precipitation grid network covering the HC and GL regions, respectively, as shown in the blue squares on the right side of Figures 3 and 4. According to the spatial distribution of precipitation observed by the GPM satellite and the actual precipitation observations from the ground meteorological stations, combined with the method mentioned in Section 2.1, a gridding data set covering the spatial distribution of precipitation in the HC and GL area can be generated, respectively.

Figure 3. Precipitation data grid structure in HC. The orange dots on the left represent the actual 8 ground meteorological stations in HC, and the blue squares on the right represent the precipitation grid data points (a total of 34) formed by dividing HC into a $0.3^\circ \times 0.3^\circ$ (latitude and longitude) grid.

Figure 4. Precipitation data grid structure in GL. The orange dots on the left represent the actual 12 ground meteorological stations in GL, and the blue squares on the right represent the precipitation grid data points (a total of 28) formed by dividing GL into a $0.3^\circ \times 0.3^\circ$ (latitude and longitude) grid.

3.3. Calculation of the Lag Time of Daily Generating Capacity

In Section 3.2, a precipitation grid data set covering HC City at 34 points and a precipitation grid data set covering GL City at 28 points are generated. The total daily precipitation in the two regions on the day and the previous 15 days (the sum of the daily precipitation values of 34 points) is taken as the input variable set to be selected, and the increment of electricity generation on the day and the previous day is taken as the output variable set to be selected, denoted as $\{P_t^s, P_{t-1}^s, P_{t-2}^s, P_{t-3}^s, \dots, P_{t-15}^s\}$ and $\{\Delta E_t\}$, respectively, where $t-i$ represents the i -th day before the date t . Using Algorithm 1 in Section 2.2 to calculate the correlation between the input variable set to be selected and the output variable set to be selected, the change in the AIC value during the calculation process is shown in Figures 5 and 6, respectively. It can be seen from the figure that the AIC value of HC city continues to decline and reaches the minimum when the 6-th variable

is calculated, while the AIC value of GL city reaches the minimum when the 4-th variable is calculated.

The above results show that the most significant impact on the fluctuation of the daily power generation of the small hydropower stations in HC is the total daily precipitation set $\{P_{t-2}^s, P_{t-1}^s, P_{t-3}^s, P_{t-4}^s, P_{t-6}^s, P_{t-5}^s\}$. Additionally, the most significant influence of ΔE_t on the daily power generation fluctuation of small hydropower station groups in GL is the total daily precipitation set $\{P_{t-1}^s, P_{t-2}^s, P_{t-4}^s, P_{t-3}^s\}$. Therefore, in the subsequent forecast of the daily power generation of the group of small hydropower stations in HC and GL, the precipitation grid data of the 6 days and 4 days before the forecast date will be used as the forecast input variable, respectively.

Figure 5. Changes in the AIC value during the lag-time selection process in HC city.

Figure 6. Changes in the AIC value during the lag-time selection process in GL city.

3.4. Filtering the Power Generation Data

Most of the small hydropower stations basically have no regulation capacity, and their power generation capacity depends on the runoff of the river. However, small hydropower stations are isolated and scattered, and it is difficult to know the relevant information of the rivers in which they are located. Therefore, it is impossible to intuitively evaluate the current level of the power generation capacity of the group of stations through the flow. We use the Gaussian weighted moving average filtering method to smooth the historical power generation data, filter out the frequent fluctuation information caused by precipitation, and obtain the general trend of recent changes in daily power generation. We also select the filtered values of the power generation on the 6 days and 4 days before the forecast date, respectively, as the assessment and prediction of the general trend of

the daily power generation change of the small hydropower station group in HC city and GL city.

3.5. Vectorization of Sample Data

Before model training, the sample data should be vectorized, and the corresponding label value should be set for each group of feature vectors.

The mathematical explanation of the forecasting problem of daily power generation of a group of small hydropower stations is to obtain a mapping function f that satisfies $f(\mathbf{X}) = Y$. Suppose the power generation in HC to be forecasted on day t is E_t ; then, $Y = E_t$. According to the analysis in Sections 3.3 and 3.4, precipitation and power generation in the 6 days before the forecasted date t should be selected as predictors, which are denoted as:

$$\mathbf{X}_1 = [PS_{t-1}, PS_{t-2}, PS_{t-3}, PS_{t-4}, PS_{t-5}, PS_{t-6}] \quad (9)$$

$$\mathbf{X}_2 = [E_{t-1}, E_{t-2}, E_{t-3}, E_{t-4}, E_{t-5}, E_{t-6}], \quad (10)$$

where $t - m$ ($m = 1, 2, 3, 4, 5, 6$) represents the m -th day before the forecasted date t . PS_{t-m} represents the spatial distribution of precipitation on the previous m -th day (its structure corresponds to the precipitation grid on the right side of Figure 3):

$$PS_{t-m} = \begin{bmatrix} 0 & P_{m,1} & P_{m,2} & P_{m,3} & P_{m,4} & P_{m,5} & P_{m,6} & 0 \\ P_{m,7} & P_{m,8} & P_{m,9} & P_{m,10} & P_{m,11} & P_{m,12} & P_{m,13} & P_{m,14} \\ 0 & P_{m,15} & P_{m,16} & P_{m,17} & P_{m,18} & P_{m,19} & P_{m,20} & P_{m,21} \\ 0 & P_{m,22} & P_{m,23} & P_{m,24} & P_{m,25} & P_{m,26} & P_{m,27} & P_{m,28} \\ 0 & P_{m,29} & P_{m,30} & P_{m,31} & P_{m,32} & P_{m,33} & 0 & 0 \\ 0 & 0 & 0 & 0 & P_{m,34} & 0 & 0 & 0 \end{bmatrix} \quad (11)$$

In Equation (11), $P_{m,j}$ ($j = 1, 2, 3, \dots, 34$) represents the precipitation value of the j -th precipitation node (corresponding to the precipitation grid point on the right side of Figure 3). In addition, 0 means that the point is not in the HC region, so the influence of its precipitation value is not considered. Finally, the input feature vector and label are:

$$\begin{cases} \mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2\} \\ Y = E_t \end{cases} \quad (12)$$

Similarly, we can also obtain the vectorized sample data of GL city.

4. Calculation Results and Discussion

4.1. Evaluation Metrics

To evaluate the forecasting method more intuitively and effectively, we choose four indicators, including accuracy (AC), mean absolute percentage error (MAPE), root mean square error (RMSE), and goodness of fit (R^2), as the evaluation basis [24]. The expressions are:

$$AC = \left[1 - \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{E_{Ri} - E_{Fi}}{E_{Ri}} \right)^2} \right] \times 100\% \quad (13)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{E_{Ri} - E_{Fi}}{E_{Ri}} \right| \times 100\% \quad (14)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (E_{Ri} - E_{Fi})^2} \quad (15)$$

$$\begin{cases} R^2 = 1 - \frac{\sum_{i=1}^n (E_{Ri} - E_{Fi})^2}{\sum_{i=1}^n (E_{Ri} - \bar{E}_R)^2} \\ \bar{E}_R = \frac{1}{n} \sum_{i=1}^n E_{Ri} \end{cases}, \quad (16)$$

where E_{Ri} and E_{Fi} are the true and forecasted daily power generation values, respectively, and n is the total number of days used for testing the model results.

4.2. Validity Analysis Considering the Spatial Distribution of Precipitation

To verify the effectiveness of considering the spatial distribution of precipitation, the forecasting results of the proposed CM-MDLN model before and after considering the spatial distribution of precipitation are compared. We divide the data into two categories. The first category considers the spatial distribution of precipitation, that is, the grid precipitation grid data generated by combining satellite remote sensing data in Section 3.2. The second category does not consider the spatial distribution of precipitation; that is, only the precipitation data observed by the original meteorological station is used. In this subsection, we take the data from June to August 2018 as an example for verification. This period belongs to the rainy season in HC and GL, and the spatial distribution of precipitation is highly variable. Therefore, this period, which is obviously representative, is selected as an example to verify the impact of the spatial distribution of precipitation on the forecast of power generation.

Figure 7 shows the comparison of the average daily precipitation in HC city from June to August and the forecasting results of daily power generation before and after considering the spatial distribution of precipitation in the model. Table 2 shows the statistical results of the evaluation metrics of the forecasting results in HC city in the two cases. Parts of the peak and valley values in the curves of the forecasting results are enlarged and displayed in Figures 8 and 9, respectively. Combining the table of statistical evaluation metrics and the magnified view of local peak and valley values, it can be seen that the fitting effect of the prediction results considering the spatial distribution of precipitation (green line, accuracy of 94.52%) is better than that of the prediction results without considering the spatial distribution of precipitation (purple line, accuracy of 88.72%) in both the trend of the curve and the peak and valley values.

Figure 7. Comparison of forecasting results before and after considering the spatial distribution of precipitation in HC city. The x-axis represents the time in days, the primary y-axis represents the power generation, and the secondary y-axis represents precipitation.

Figure 8. Magnified view of local peak values of forecasting results of HC city. The x -axis represents the time in days, and the y -axis represents the power generation.

Figure 9. Magnified view of local valley values of forecasting results of HC city. The x -axis represents the time in days, and the y -axis represents the power generation.

Table 2. Statistics of evaluation metrics for forecasting results before and after considering the spatial distribution of precipitation in HC city.

Evaluation Metrics	Is the Spatial Distribution of Precipitation Considered?	
	No	Yes
AC (%)	88.72	94.52
MAPE (%)	8.87	4.55
RMSE (MWh)	570.26	288.43
R^2 ($\times 10^{-1}$)	5.385	8.564
$APE < 10\%$ (d) ¹	58	86
$APE < 10\%$ (%) ²	63.04	93.48
$APE < 5\%$ (d) ³	31	57
$APE < 5\%$ (%) ⁴	33.70	61.96

¹ The number of days with APE less than 10%. ² The percentage of days with APE less than 10%. ³ The number of days with APE less than 5%. ⁴ The percentage of days with APE less than 5%.

Figure 10 shows the comparison of the average daily precipitation in HC city from June to August and the forecasting results of daily power generation before and after considering the spatial distribution of precipitation in the model. Figures 11 and 12 show the enlarged display of some peaking and valley values in the curve of the forecasting results. Table 3 shows the statistical results of evaluation metrics of the forecasting results in GL under two conditions. Combining the results in Figures 10–12 and Table 3, it can also be clearly

seen that the prediction results after considering the spatial distribution of precipitation are better than those without considering the spatial distribution of precipitation.

Figure 10. Comparison of forecasting results before and after considering the spatial distribution of precipitation in GL city.

Figure 11. Magnified view of local peak values of forecasting results of HC city.

Figure 12. Magnified view of local peak values of forecasting results of GL city.

Table 3. Statistics of evaluation metrics for forecasting results before and after considering the spatial distribution of precipitation in GL city.

Evaluation Metrics	Is the Spatial Distribution of Precipitation Considered?	
	No	Yes
AC (%)	85.02	93.39
MAPE (%)	11.77	5.63
RMSE (MWh)	528.66	217.15
$R^2 (\times 10^{-1})$	8.347	9.721
$APE < 10\%$ (d) ¹	47	82
$APE < 10\%$ (%) ²	51.09	89.13
$APE < 5\%$ (d) ³	25	43
$APE < 5\%$ (%) ⁴	27.17	46.74

¹ The number of days with APE less than 10%. ² The percentage of days with APE less than 10%. ³ The number of days with APE less than 5%. ⁴ The percentage of days with APE less than 5%.

4.3. Comparison of Forecasting Models

To verify the effectiveness and universality of the proposed CM-MDLN forecasting method, we compare the forecasting results of different methods based on the same data set mentioned above in this subsection. The following six methods are compared: support vector regression (SVR), gradient boosting regression tree (GBRT), random forest (RF), long-short-term memory network (LSTM), and, separately, MLP and CNN. The above comparison methods are all based on the Keras deep learning framework in the Python development environment. The SVR, GBRT, and RF models are implemented by calling the Sklearn machine learning library. The SVR model takes the radial basis function (RBF) as the kernel function, and the penalty coefficient C is set to 10. The GBRT and RF models take default parameter values. The LSTM model takes a three-layer network, with 64, 32, and 16 neurons in each of the three layers. The separate MLP and CNN models have the same network parameter settings as the proposed model. Since several methods such as SVR, GBRT, RF, LSTM, and separate MLP cannot directly process the spatial distribution data of precipitation, it is necessary to reduce the two-dimensional spatial distribution data of precipitation to the one-dimensional vector data before using these methods.

To verify the universality of the proposed method, the HC data set is randomly divided into a training set, validation set, and testing set according to proportions of 80%, 10%, and 10%, respectively. Since the proportion of non-precipitation periods in a year is higher than that of precipitation periods, the data of the precipitation periods and non-precipitation periods contained in the randomly divided data sets are relatively evenly distributed, and the calculated results are more convincing. Among them, the training set is used for model training, the validation set is used for tuning model parameters, and the testing set is used to test the forecasting effect of the model.

The comparison of the forecasted value and the true value of each model in HC city and GL city are shown in Figures 13 and 14 (the three straight lines in each figure represent $y = 1.1x$, $y = x$, and $y = 0.9x$ from top to bottom). In the results of HC city and GL city, almost all the points of the proposed model fall between the two lines $y = (1 \pm 0.1)x$ and are closer to the diagonal line $y = x$, which indicates that the forecasted values of this model are closer to the true values and that the forecasting accuracy is higher than that of other models.

Figure 13. Scatter plots for seven methods in HC city: (a) SVR; (b) GBRT; (c) RF; (d) LSTM; (e) MLP; (f) CNN; (g) proposed CM-MDLN method. The x -axis represents the true value, and the y -axis represents the forecasted value corresponding to the true value. The three straight lines represent $y = 1.1x$, $y = x$, and $y = 0.9x$ from top to bottom.

Figure 14. Scatter plots for seven methods in GL city: (a) SVR; (b) GBRT; (c) RF; (d) LSTM; (e) MLP; (f) CNN; (g) proposed CM-MDLN method.

The evaluation metric statistics of the 106-day power generation forecast results in the testing set of HC city and GL city are shown in Tables 4 and 5, and the comparison between the forecasted values and the true values of each model is shown in Figure 8. It is easy to see that the *AC*, *MAPE*, *RMSE*, and R^2 of the proposed multimodal deep learning model based on the fusion of a CNN and an MLP are significantly better than those of the SVR, GBRT, RF, LSTM, and separate MLP and CNN models. In particular, the percentage of days with absolute percentage error (*APE*) less than 10% or 5% is much higher than other models. This proves that the proposed model is more effective and universal, and it is more suitable for forecasting the daily power generation of practical small hydropower stations.

Table 4. Statistics of evaluation metrics for forecasting results of different methods in the HC city testing set.

Evaluation Metrics	Methods						
	SVR	GBRT	RF	LSTM	MLP	CNN	CM-MDLN
<i>AC</i> (%)	87.31	86.77	87.97	84.91	86.32	85.87	93.07
<i>MAPE</i> (%)	8.70	9.01	8.30	10.72	9.92	9.16	5.71
<i>RMSE</i> (MWh)	326.28	277.51	267.66	418.31	313.40	279.93	238.41
R^2 ($\times 10^{-1}$)	9.613	9.692	9.607	9.504	9.635	9.689	9.847
<i>APE</i> < 10% (d) ¹	74	80	78	64	70	74	92
<i>APE</i> < 10% (%) ²	69.81	75.47	73.58	60.38	66.04	69.81	86.79
<i>APE</i> < 5% (d) ³	44	46	47	37	30	45	55
<i>APE</i> < 5% (%) ⁴	41.51	43.39	44.34	34.91	28.30	42.45	51.89

¹ The number of days with *APE* less than 10%. ² The percentage of days with *APE* less than 10%. ³ The number of days with *APE* less than 5%. ⁴ The percentage of days with *APE* less than 5%.

Table 5. Statistics of evaluation metrics for forecasting results of different methods in the GL city testing set.

Evaluation Metrics	Methods						
	SVR	GBRT	RF	LSTM	MLP	CNN	CM-MDLN
<i>AC</i> (%)	88.70	89.60	89.72	84.78	85.79	86.51	92.80
<i>MAPE</i> (%)	7.45	7.52	7.40	11.59	9.57	10.33	5.70
<i>RMSE</i> (MWh)	305.29	290.99	311.98	429.62	344.93	362.73	210.63
R^2 ($\times 10^{-1}$)	9.714	9.740	9.702	9.435	9.635	9.597	9.808
<i>APE</i> < 10% (d) ¹	80	79	79	60	70	64	82
<i>APE</i> < 10% (%) ²	75.47	74.52	74.52	56.60	66.03	60.37	77.36
<i>APE</i> < 5% (d) ³	51	47	52	33	43	35	56
<i>APE</i> < 5% (%) ⁴	48.11	44.33	49.05	31.13	40.56	33.01	52.83

¹ The number of days with *APE* less than 10%. ² The percentage of days with *APE* less than 10%. ³ The number of days with *APE* less than 5%. ⁴ The percentage of days with *APE* less than 5%.

5. Conclusions

This paper proposes a multimodal deep learning method for forecasting the daily power generation of small hydropower stations considering the temporal and spatial distribution of precipitation. The characteristics of the method are as follows:

- (1) Precipitation grid data with spatial distribution differences are applied to the neural network model to deeply explore the influence of the spatial distribution of precipitation on the daily power generation of a group of small hydropower stations.
- (2) The time lag between daily power generation changes and precipitation is analyzed, and the PMI method is used to estimate this “time difference”, which is used to select the best time-scale precipitation data for forecasting the daily power generation of a group of small hydropower stations.
- (3) Using multimodal deep learning methods based on a CNN and an MLP, according to the different characteristics of each modal data, different methods can be used to fully extract and integrate the characteristic information hidden in the precipitation and historical power generation data, improving the accuracy of forecasts of the daily power generation of a group of small hydropower stations.

Through the simulation analysis of real data in southern China, the results show that the proposed method takes into account the spatial distribution of precipitation and can effectively improve the accuracy and other evaluation metrics of the forecasts of the daily power generation of a group of small hydropower stations. It can provide a basis for the main power network in this region to solve the problem of “blind adjustment” of small hydropower generation, provide security for the power network, and provide decision-making support for the energy balance of the local power dispatching department. Finally, it can reduce the abandoned water of small hydropower stations and increase clean and renewable power generation.

Author Contributions: S.Y. proposed and conceptualized the framework of the method and wrote the full text. H.W. supervised and reviewed the research. L.Z. and S.Q. have provided a lot of advice on the modeling process and paper writing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (NSFC), grant number 51967002.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from S.Y.

Acknowledgments: Many thanks are offered to the National Natural Science Foundation of China (Grant No. 51967002).

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

References

1. United Nations Industrial Development Organization. World Small Hydropower Development Report 2019. Available online: <https://www.unido.org/sites/default/files/files/2020-07/ASIA%20PACIFIC%20Book.pdf> (accessed on 3 July 2021).
2. Department of Rural Water Resources and Hydropower, Ministry of Water Resources of the People's Republic of China. 2019 Annual Report on Rural Water Conservancy and Hydropower Work. Available online: http://www.mwr.gov.cn/sj/tjgb/ncslsdbn/202007/t20200730_1430330.html (accessed on 6 April 2021).
3. Kong, Y.; Wang, J.; Kong, Z.; Song, F.; Liu, Z.; Wei, C. Small hydropower in China: The survey and sustainable future. *Renew. Sustain. Energy Rev.* **2015**, *48*, 425–433. [\[CrossRef\]](#)
4. Cheng, C.; Liu, B.; Chau, K.-W.; Li, G.; Liao, S. China's small hydropower and its dispatching management. *Renew. Sustain. Energy Rev.* **2015**, *42*, 43–55. [\[CrossRef\]](#)
5. Magnusson, J.; Nævdal, G.; Matt, F.; Burkhart, J.F.; Winstral, A. Improving hydropower inflow forecasts by assimilating snow data. *Hydrol. Res.* **2020**, *51*, 226–237. [\[CrossRef\]](#)
6. Reichl, F.; Hack, J. Derivation of flow duration curves to estimate hydropower generation potential in data-scarce regions. *Water* **2017**, *9*, 572. [\[CrossRef\]](#)
7. Jingsong, D.; Difei, S.; Hanting, Y.; Zhaolong, W. Stochastic modeling of small HydroPower output based on regional synchronism feature and its allowed penetration level research. In Proceedings of the 2018 International Conference on Power System Technology (POWERCON), Guangzhou, China, 6–8 November 2018; pp. 1433–1438.
8. Wen, X.; Gao, X.; Su, L.; Fan, Q.; Peng, W. Analysis of influencing factors and characteristics of small hydropower generation. In Proceedings of the 2017 9th International Conference on Modelling, Identification and Control (ICMIC), Kunming, China, 10–12 July 2017; pp. 658–662.
9. Li, G.; Li, B.J.; Yu, X.G.; Cheng, C.T. Echo state network with Bayesian regularization for forecasting short-term power production of small hydropower plants. *Energies* **2015**, *8*, 12228–12241. [\[CrossRef\]](#)
10. Li, G.; Liu, C.X.; Liao, S.L.; Cheng, C.T. Applying a correlation analysis method to long-term forecasting of power production at small hydropower plants. *Water* **2015**, *7*, 4806–4820. [\[CrossRef\]](#)
11. Cheng, C.T.; Miao, S.M.; Luo, B.; Sun, Y.J. Forecasting monthly energy production of small hydropower plants in ungauged basins using grey model and improved seasonal index. *J. Hydroinform.* **2017**, *19*, 993–1008. [\[CrossRef\]](#)
12. Kim, C.; Kim, D.-H. Effect of rainfall spatial distribution and duration on minimum spatial resolution of rainfall data for accurate surface runoff prediction. *J. Hydro-Environ. Res.* **2018**, *20*, 1–8. [\[CrossRef\]](#)
13. Meng, X.Y.; Wang, H.; Shi, C.X.; Wu, Y.P.; Ji, X.N. Establishment and evaluation of the China meteorological assimilation driving datasets for the SWAT model (CMADS). *Water* **2018**, *10*, 1555. [\[CrossRef\]](#)
14. Di, Z.; Maggioni, V.; Mei, Y.; Vazquez, M.; Houser, P.; Emelianenko, M. Centroidal Voronoi tessellation based methods for optimal rain gauge location prediction. *J. Hydrol.* **2020**, *584*, 124651. [\[CrossRef\]](#)

15. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; Chapman & Hall: London, UK, 1986; Volume 39, pp. 296–297.
16. May, R.J.; Maier, H.R.; Dandy, G.C.; Fernando, T.M.K.G. Non-linear variable selection for artificial neural networks using partial mutual information. *Environ. Model. Softw.* **2008**, *23*, 1312–1326. [[CrossRef](#)]
17. Hanchuan, P.; Fuhui, L.; Ding, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1226–1238. [[CrossRef](#)] [[PubMed](#)]
18. Sharma, A.; Luk, K.C.; Cordery, I.; Lall, U. Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 2—Predictor identification of quarterly rainfall using ocean-atmosphere information. *J. Hydrol.* **2000**, *239*, 240–248. [[CrossRef](#)]
19. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [[CrossRef](#)]
20. Baltrušaitis, T.; Ahuja, C.; Morency, L. Multimodal machine learning: A survey and taxonomy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 423–443. [[CrossRef](#)] [[PubMed](#)]
21. Ramachandram, D.; Taylor, G.W. Deep multimodal learning: A survey on recent advances and trends. *IEEE Signal Process. Mag.* **2017**, *34*, 96–108. [[CrossRef](#)]
22. McGurk, H.; MacDonald, J. Hearing lips and seeing voices. *Nature* **1976**, *264*, 746–748. [[CrossRef](#)] [[PubMed](#)]
23. NASA. The Global Precipitation Measurement Mission (GPM). Available online: <https://gpm.nasa.gov/missions/GPM#gpmcoreobservatorysatellite> (accessed on 20 March 2021).
24. Kim, S.; Kim, H. A new metric of absolute percentage error for intermittent demand forecasts. *Int. J. Forecast.* **2016**, *32*, 669–679. [[CrossRef](#)]