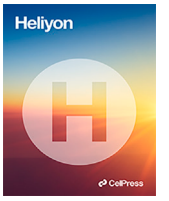
Heliyon 10 (2024) e28312

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

# Optimizing hydropower scheduling through accurate power load prediction: A practical case study

Guangqin Huang [a], Ming Tan [a], Zhihang Meng [b,c], Jiaqi Yan [d], Jin Chen [e], Qiang Qu [c,*]

[a] Guizhou Wujiang River Navigation Authority, Tongren, 565100, Guizhou, China
[b] College of Mathematics and Information Science, Hebei University, Baoding, 071002, Hebei, China
[c] Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, 518055, Guangdong, China
[d] Guizhou Silin Navigation Authority, Tongren, 565100, Guizhou, China
[e] Guizhou Zhongnan Transport Technology co. Ltd., Guiyang, 550000, Guizhou, China

ARTICLE INFO

*MSC:*
0000
1111

*Keywords:*
Multi-objective optimization
Scheduling strategy
Hydropower station
Prediction algorithm
Neural networks
Deep learning

ABSTRACT

Hydropower stations that are part of the grid system frequently encounter challenges related to the uneven distribution of power generation and associated benefits, primarily stemming from delays in obtaining timely load data. This research addresses this issue by developing a scheduling model that combines power load prediction and dual-objective optimization. The practical application of this model is demonstrated in a real-case scenario, focusing on the Shatuo Hydropower Station in China. In contrast to current models, the suggested model can achieve optimal dispatch for grid-connected hydropower stations even when power load data is unavailable. Initially, the model assesses various prediction models for estimating power load and subsequently incorporates the predictions into the GA-NSGA-II algorithm, specifically an enhanced elite non-dominated sorting genetic algorithm. This integration is performed while considering the proposed objective functions to optimize the discharge flow of the hydropower station. The outcomes reveal that the CNN-GRU model, denoting Convolutional Neural Network-Gated Recursive Unit, exhibits the highest prediction accuracy, achieving R-squared and RMSE (i.e., Root Mean Square Error) values of 0.991 and 0.026, respectively. The variance between scheduling based on predicted load values and actual load values is minimal, staying within 5 ($m^3/s$), showcasing practical effectiveness. The optimized scheduling outcomes in the real case study yield dual advantages, meeting both the demands of ship navigation and hydropower generation, thus achieving a harmonious balance between the two requirements. This approach addresses the real-world challenges associated with delayed load data collection and insufficient scheduling, offering an efficient solution for managing hydropower station scheduling to meet both power generation and navigation needs.

## 1. Introduction

The development of water conservancy projects exerts a significant influence on the agricultural and economic progress of contemporary society, serving as a pivotal factor in strategic, foundational, and socially advantageous aspects. As a result, there is a growing focus on improving the operational effectiveness of hydropower stations that are interconnected with the power grid.

* Corresponding author.
*E-mail address:* qiang@siat.ac.cn (Q. Qu).

https://doi.org/10.1016/j.heliyon.2024.e28312
Received 18 June 2023; Received in revised form 11 March 2024; Accepted 15 March 2024

Available online 21 March 2024

In river basins with navigation requirements, the coexistence of power generation and navigation introduces mutual constraints, posing contradictory scheduling objectives. Lack of equilibrium between navigation requirements and power generation substantially reduces the operational effectiveness of hydropower stations. Efficiently managing the scheduling of hydropower stations, taking into account the needs of both hydropower generation and ship navigation, poses a significant challenge within industrial systems. This task is crucial for aligning the objectives of ship navigation and power generation, contributing to the ongoing endeavors in attaining carbon neutrality [1].

Yet, difficulties emerge during the practical scheduling phase, especially when it comes to obtaining grid power load data promptly. Obtaining grid power load data for hydropower stations turns out to be a intricate undertaking. In the initial stages, the grid operator coordinates power market transactions to ascertain the power demand, following which the hydropower station collaborates with the grid operator to formulate its hydropower generation plan. In this procedure, encountering delays in obtaining power load data promptly is a frequent issue. The failure to swiftly and efficiently acquire load demand curves directly affects the optimal scheduling outcomes for hydropower stations, impacting discharge flow, water levels, and the overall efficacy of ship navigation. It is imperative to promptly devise practical and effective methods to tackle this challenge.

To address such challenges, the initial step involves formulating these issues into a mathematical model with constraints, as outlined by Tirkolaee et al. [2]. However, existing research has predominantly concentrated on the benefits for hydropower stations, navigation, and ecological impacts. Notably, there has been a lack of systematic attention to the scientific and optimal dispatch of hydropower stations, particularly in scenarios where power load data are not promptly available. Confronted with a power load shortage, current optimal scheduling models are constrained to utilize historical data with the aim of achieving scheduling for maximum efficiency.

Therefore, this research presents a unified scheduling framework for both hydropower generation and ship navigation. The framework effectively combines prediction and optimization algorithms to manage the scheduling of a hydropower station in instances where timely grid power load data are not accessible. It collects discharge flow data for each time interval, meeting the requirements of both grid power generation and ship navigation. Contrasting with prior literature, the contribution of this study can be summarized as follows:

1. Integrated Scheduling Model: A unique dispatching model is introduced, utilizing power load prediction and dual-objective optimization within an actual station system. This model creates a cohesive scheduling framework that integrates prediction with subsequent scheduling, ensuring the hydropower station's operations are both scientific and accurate in situations where power load data is unavailable. The model improves the collaborative efficiency of the scheduling process.
2. Hydropower dual-objective Optimization: Introducing dual-objective functions considering practical constraints of hydropower stations, including water level, flow rate, and unit output. The functions put emphasis on enhancing power generation efficiency by targeting the minimum power generation minus load demand. This results in the formulation of an optimization process and its corresponding solution for achieving the optimal scheduling of hydropower stations.
3. Practical Implementation and Case Study: The proposed model is deployed in the Shatuo Hydropower Station system, and its performance is assessed in real-world applications. This practical case study serves to deepen comprehension and demonstrate the practical applicability of hydropower station scheduling, particularly in situations where power load data is unavailable.

The rest of the paper is structured as follows: Section 2 provides a brief overview of relevant studies, Section 3 delves into the models to predict power loads, Section 4 outlines the proposed solutions of the joint scheduling model, Section 5 presents the performance evaluation, and finally, Section 6 wraps up the paper with considerations for future work.

## 2. Related work

Grid power load prediction analyzes the evolving patterns in historical power load data to forecast future load accurately. In the context of power grid system development, strategically organizing grid power supply loads is vital for hydropower station stability and efficiency. Numerous effective methods have been proposed by researchers to enhance the precision of power load prediction. For instance, traditional time series models such as Auto Regressive (AR), Moving Average (MA), and AR Integrated MA (ARIMA) are widely employed for this task. Alberg et al. utilized the sliding window-based ARIMA algorithm to forecast power load, assessing its performance against several machine learning algorithms, including Random Forest (RF) [3]. Chen et al. employed the Support Vector Regression (SVR) algorithm, a kind of the Support Vector Machine (SVM) algorithm recognized for its simplicity, in their short-term power load prediction study [4]. Xu et al. [5] employed the Long Short-Term Memory network (LSTM), utilizing the extended memory capacity to forecast power load sequences based on long-term dependencies. Both the Gate Recurrent Unit (GRU) model and LSTM, which represent advancements in the RNN algorithm, have shown success. Liu et al. [6] applied the GRU model for predicting energy consumption, surpassing the performance of the Multiple Linear Regression (MLR) model. The Convolutional Neural Network (CNN) model, originally designed for images, has been adapted for predicting time series. Tudose et al. [7] applied the CNN model to predict power load using a publicly available dataset from the Romanian power system. In recent times, researchers have integrated models to improve the accuracy. Cheng et al. [8] introduced an enhanced CNN-LSTM model for the task of predicting power load data, considering seasonal effects. Sajjad et al. [9] presented a hybrid CNN-GRU model designed for predicting power consumption, assessing its effectiveness across various datasets. These machine learning and deep learning algorithms have demonstrated success in predicting power load. To attain the optimal prediction from the case study datasets, we

chose seven of these models for experimentation and identified the most suitable option for real-world hydropower stations through comparison.

Scheduling hydropower stations poses a multi-objective optimization problem. Currently, there exist various algorithms designed to address multi-objective optimization, each with its distinct advantages and drawbacks. Genetic Algorithm (GA) [10] is well-suited for both single and multi-objective optimization, boasting strong global search capabilities and robustness. However, modifications are necessary when dealing with multi-objective problems to ensure the preservation of a set of non-dominated solutions. The Nondominated Sorting Genetic Algorithm (NSGA) demonstrates rapid convergence and effective distribution for simultaneous optimization of multiple objectives. Yet, it requires managing a substantial number of non-dominated solutions and is sensitive to parameter selection. NSGA-II [11,12] surpasses its predecessor in performance and convergence, proving more effective in maintaining sets of non-dominated solutions. Nevertheless, challenges persist in high-dimensional problems and parameter selection. The Multi-Objective Particle Swarm Optimization (MOPSO) [13] utilizes moving particles to explore optimal solutions for multiple objectives, applicable to continuous, discrete, and mixed optimization problems. However, it may encounter limitations in high-dimensional problems due to the curse of dimensionality. Multi-Objective Genetic Algorithm (MOGA) [14] maintains a set of non-dominated solutions through genetic operations, offering multiple excellent solutions for selection across various optimization problems. Nevertheless, it may encounter challenges in searching for solutions in high-dimensional problems and demands substantial storage space for maintaining the solution set.

In recent research, considerable attention has been given to multi-objective optimization, especially in the context of hydropower stations. For instance, Liu et al. [15] created a model that considers power generation, output stability, and downstream river ecology, employing an improved multi-objective optimization evolutionary algorithm. Meng et al. [16] presented an enhanced multi-objective cuckoo search algorithm (IMOCS) to optimize the balance between water and energy at the Xiaolangdi Hydropower Station on the Yellow River. This algorithm, validated using five indicators, exhibited successful outcomes. Meanwhile, Yang et al. [17] applied the shuffled frog leaping algorithm (SFLA) to address the multi-objective optimization challenge involving power generation and ecological objectives in cascade reservoirs, termed the multi-objective ecological operation for cascade reservoirs (MOEOCR). Fang et al. [18] tackled the coordination of power generation benefits and ecological protection at the Minjiang Shuikou Hydropower Station in China. They employed an enhanced multi-objective particle swarm optimization algorithm (MOPSO) in conjunction with the self-organizing mapping (SOM) method. In a similar vein, Qin et al. [19] utilized the Relative Speed Estimation Algorithm (RSEA) scheduling algorithm to optimize the scheduling of the Shuibuya Hydropower Station, considering power generation, navigation, and ecological objectives. They compared the results with a classical scheme. Additionally, Yu et al. [20] investigated the influence of climate and hydrology on hydropower station performance. They formulated a multi-objective optimization problem that considered the total power output and remaining load changes in the grid for the Jinsha River. Kong et al. [21] presented a marine predators algorithm designed to maximize power generation while minimizing adverse ecological effects. They conducted a performance comparison with other algorithms. In a related context, Chen et al. [22] applied three multi-objective optimization algorithms to formulate a coordinated model for power generation and ecological flow in a terraced hydropower station. They compared the outcomes of these three models. Additionally, Zhou et al. [23] introduced a scheduling model that takes into account the energy storage mechanism. The predominant emphasis in current research lies in optimizing the synergy between water resources and ecology for hydrapower generation. Limited attention has been given to investigating methods for enhancing shipping efficiency while concurrently satisfying the power load demand of the grid. Furthermore, there is a notable gap in addressing the implications of lacking power load data on the optimal scheduling of hydropower stations.

## 3. Power load data forecasting model

Access to power load data for our selected practical case, the Shatuo Hydropower Station, is only available after 9:00 pm every night, presenting a challenge in terms of timeliness. In order to overcome this obstacle, we utilize historical grid power load data to predict the future needs of the grid.

To address the challenge in this practical scenario, we conduct a comparative study on existing prediction models to identify the most suitable option for hydropower stations. Specifically, we compare seven prediction algorithms, including widely used and popular models such as Support Vector Regression (SVR) [24], Auto Regressive (AR), Moving Average Integrated Moving Average (ARIMA) [25], Convolutional Neural Network (CNN) [26], Long Short-Term Memory network (LSTM) [27], Gate Recurrent Unit (GRU) [28], CNN-LSTM [29], and CNN-GRU [9]. Ultimately, the results lead us to adopt CNN-GRU with minor adjustments for the prediction task in practice.

The Convolutional Gated Recurrent Unit (CNN-GRU) is a hybrid model that synergistically incorporates the strengths of Convolutional Neural Networks (CNN) and Gated Recurrent Units (GRU), offering robust expressive power and superior generalization capabilities. This model excels in processing sequential data and extracting features. To begin with, CNN is deployed to extract feature vectors, serving as inputs to the subsequent GRU layer. The GRU neural network effectively captures dynamic patterns within these features, facilitating accurate predictions. This approach empowers the model to efficiently mine information from temporal data, thereby enhancing prediction accuracy.

The power load data originating from the Shatuo hydropower station is presented as a time-series dataset, wherein past power load information influences the current demand at each moment. Furthermore, during the practical scheduling process, the power load data displays variability, and the length of each input sequence is not constant, necessitating adaptable adjustments based on real-time inputs. The CNN-GRU model, with its ability to capture contextual information in time series data and its flexibility with variable-length temporal data, emerges as an optimal selection for predicting the power load of the station.

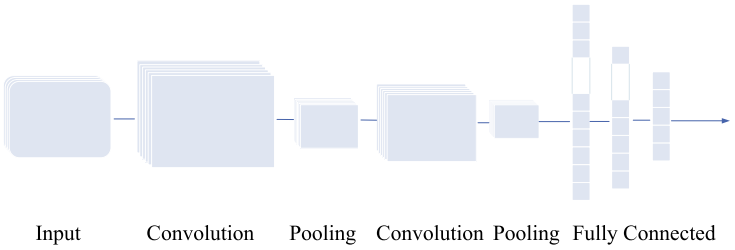



**Fig. 1.** CNN illustration.

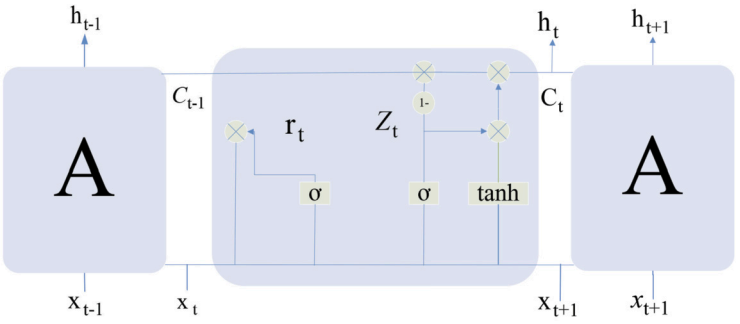



**Fig. 2.** GRU illustration.

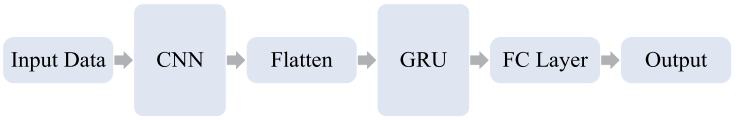



**Fig. 3.** CNN-GRU illustration.

The CNN-GRU temporal prediction model starts by inputting a time-series dataset. The CNN model then engages in feature extraction, treating the input data from each moment as an input. Following this, the features generated by the CNN are flattened into a vector, which serves as input to the GRU model. In the GRU layer, the flattened features are input, and the GRU computes the current hidden state $H(t)$ at each time step based on the current input $F(t)$ and the hidden state $H(t-1)$ from the previous time step. Optionally, a fully-connected layer can be introduced after the GRU layer for additional processing of the GRU output. Lastly, the output layer is utilized to produce the timing prediction result. The detailed structures are depicted in Figs. 1-3, respectively [9, 26,28].

## 4. The model for coordinated scheduling of power generation and navigation

The efficiency of power generation and navigation at the Shatuo Hydropower Station is closely intertwined with its downstream flow, giving rise to a mutual constraint and contradiction between these two objectives. This situation falls within the domain of a Pareto problem in multi-objective optimization [30]. To address this challenge, the study commences by scrutinizing the factors influencing navigation efficiency and examining the correlation between discharge flow and the navigation assurance rate of the hydropower station. Subsequently, it investigates the inherent constraints in the practical scheduling process of the Shatuo Hydropower Station. Lastly, the study defines the objective functions for both hydrapower generation and shipping, addressing the issue using the GA-NSGA-II algorithm.

### *4.1. Investigating variables influencing shipping efficiency*

Through an analysis of aqueous elements affecting navigation, it has been established that the navigational safety of the downstream channel at the Shatuo Hydropower Station is primarily linked to fluctuations in water level, flow rate, and downstream channel slope [31]. Particularly, variations in downstream water level, daily tailwater level changes, and water flow rates exhibit a robust association with discharge flow. Changes in downstream flow, whether an increase or decrease, directly affect both downstream water levels and flow rates. Moreover, the rate of water level variation and the downstream water level are intertwined with discharge flow. Therefore, the key factor influencing downstream channel safety is the discharge flow. Specific factors affecting navigation safety are outlined in Table 1.

**Table 1**
Factors affecting safety of navigation.

| Influencing Factors | Description |
|---|---|
| Discharge flow | To meet navigational requirements, the discharge flow needs to fall within a certain range. |
| Daily tailwater level variation | The daily fluctuation of water levels is positively correlated with the discharge flow. Due to the hydropower station's responsibility for peak load regulation in the power grid, the discharge flow experiences significant fluctuations, leading to the generation of non-uniform flow in the navigation channel. Therefore, specific conditions must be met for the daily tailwater level fluctuations as well. |
| Hourly tailwater level variation | For considerations related to navigational safety, in addition to meeting the conditions for daily water level variations, there is also a requirement for hourly variations. |
| Downstream water level variation | An indicator reflecting the rate of change in downstream water levels. |
| Water flow rate | There are specific requirements for water flow velocity when ships navigate through. |
| Downstream Channel Slope | Whether ships are moving upstream or downstream, a significant channel slope can impact ship safety. |

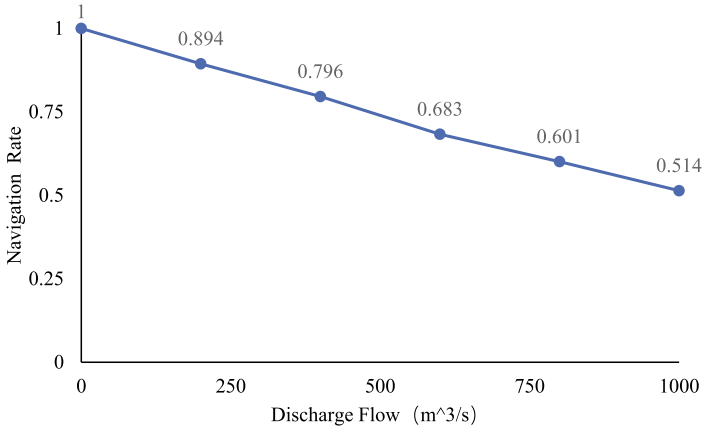



**Fig. 4.** Correlation between discharge flow and navigation rate.

Upon analyzing the historical data from the Shatuo Hydropower Station, it was observed that the navigation rate exceeds 0.9 when the discharge flow rate is in the range of 0-200 $m^3/s$, and it remains above 0.8 when the discharge flow rate is in the range of 200-400 $m^3/s$. However, as the discharge flow increases, there is a subsequent decrease in the navigation rate. The correlation between the specific discharge flow of the Shatuo Hydropower Station and the navigation rate is depicted in Fig. 4, with the relationship between the two demonstrating near linearity.

### *4.2. Practical constraints*

The scheduling system of the Shatuo Hydropower Station is subject to the following constraints.

Constraints on the operating water level of the hydropower station: The hydropower station is constrained by water level limits, which must be maintained within specified upper and lower thresholds to ensure operational stability and downstream channel safety, as shown in Equation (1):

$$Z^j_{min} \le Z^j \le Z^j_{max}. \tag{1}$$

Here, $Z^j$ denotes the operational water level of the power station at time $j$, while $Z^j_{max}$ and $Z^j_{min}$ represent the upper and lower limits of the operational water level of the power station at time $j$, respectively. The water level trend of the Shatuo Hydropower Station is illustrated in Fig. 5, where the red line signifies the upper limit of water level operation, and the blue line represents the lower limit of water level operation.

Constraints on the discharge flow from the hydropower station: The discharge flow rate is intricately linked to downstream navigational benefits and ecological stability. It must adhere to specific conditions, as shown in Equation (2):

$$Q^j_{min} \le Q^j \le Q^j_{max}. \tag{2}$$

Here, $Q^j$ signifies the discharge flow of the power station at time $j$, while $Q^j_{max}$ and $Q^j_{min}$ denote the maximum and minimum discharge flow required by the power station to meet shipping requirements at time $j$, respectively. The flow pattern of the Shatuo Hydropower Station is illustrated in Fig. 6.

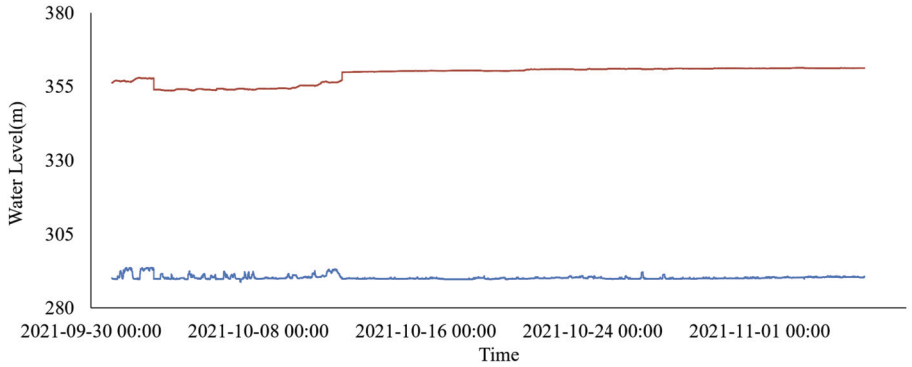



**Fig. 5.** Pattern of tailwater level and upstream water level.

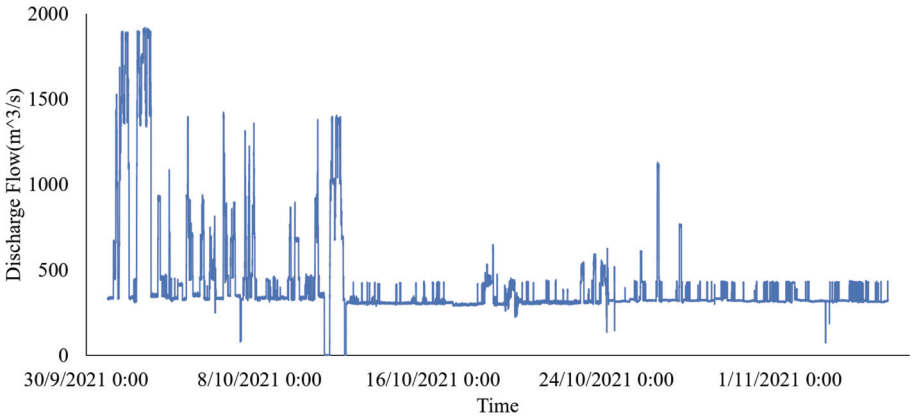



**Fig. 6.** The flow trend of the Shatuo Hydropower Station.

Constraints on the Periodic Output of the Hydropower Station: To ensure the protection of the hydropower station units and maintain the stability of peak load regulation in the power grid, the output must fall within a specified range, as shown in Equation (3).

$$P_{min}^{j} \le P^{j} \le P_{max}^{j}, \tag{3}$$

where $P_{max}^{j}$ and $P_{min}^{j}$ are the maximum and minimum allowable output of each power unit in the hydropower station, respectively.

Constraints on Downstream Water Level Variation: The downstream water level must not only adhere to upper and lower limits but also meet specified values for the rate of water level change within a period, as shown in Equations (4), (5) and (6).

$$\Delta Z_d \le Z_{d\_max}, \tag{4}$$

$$\Delta Z_h \le \Delta Z_{h\_max}, \tag{5}$$

$$Z_{15min} \le Z_{15min\_max}, \tag{6}$$

where $\Delta Z_d$, $\Delta Z_h$, $Z_{15min}$ are the daily, hourly and 15-minute variations of the downstream water level respectively. $Z_{d\_max}$, $\Delta Z_{h\_max}$, $Z_{15min\_max}$ are the corresponding maximum values.

Non-negativity Constraints: All variables mentioned above must be non-negative.

### *4.3. Objective function for the joint scheduling*

The objective of this paper is to optimize both the hydrapower generation and ship navigation scheduling goals for the station, concurrently addressing the needs of power grid peak regulation and shipping. In order to maximize water resource utilization and minimize the impact on ship navigation, two objective functions are chosen: the power generation and the shipping objectives.

Power generation objective: The goal of generation scheduling is to minimize the result obtained by subtracting the predicted demand from the overall power generation of the hydropower station, considering the initial and final water levels throughout the calculation period. This optimization goal, referred to as “power by water”, is mathematically expressed as shown in Equation (7):

$$\min f_1 = \left\{ 9.81 \times \sum_{j=1}^{T} P^j \Delta t - F_e \right\} = \left\{ 9.81 \times \sum_{j=1}^{T} H^j Q^j \Delta t - F_e \right\}. \tag{7}$$

Here, $f_1$ signifies the optimization objective for generation scheduling, representing the total power generation of the hydropower station during the time period $T$. The coefficient 9.81 corresponds to the power output coefficient of the station, while $\Delta t$ denotes the calculated time interval. $F_e$ represents the predicted grid demand. Additionally, $P^j$, $H^j$, and $Q^j$ respectively stand for the power output, hydraulic head, and generation quoted flow of the power station during the $j$th time period.

Shipping Objective: With the established mapping relationship between the discharge flow of the hydropower station and the navigation rate, the latter is employed as the evaluation metric for channel navigation efficiency. A higher value indicates superior navigation efficiency. The shipping scheduling function is as shown in Equation (8):

$$\max f_2 = \frac{1}{T} \sum_{j=1}^{T} k_j \left( Q^j \right). \tag{8}$$

In this context, $f_2$ denotes the objective function for downstream channel shipping scheduling, with $k_j$ representing the assured rate of navigation at time $j$, and $Q^j$ indicating the discharge flow of the hydropower station at time $j$.

### 4.4. GA-NSGA-II scheduling algorithm

The genetic algorithm (GA) serves as a computational simulation technique inspired by natural selection and competition principles, applied for problem-solving. The key stages of the algorithm involve encoding, establishing the fitness function, selection, crossover, and mutation, repeated iteratively until an optimal solution is attained. Unlike conventional mathematical planning algorithms, GA provides extensive applicability, a clear algorithmic structure, and straightforward parallelization processing. Nevertheless, it demonstrates constrained global search capability and vulnerability to local optimal solutions. Conversely, the non-dominated ranking genetic algorithm (NSGA-II) sets itself apart from the GA algorithm by integrating a ranking stratification rooted in dominance relationships before engaging in genetic operations. It utilizes an elite strategy to safeguard the most well-adapted individuals, maintaining overall diversity through the assessment of crowding degrees among individuals sharing comparable priority levels. In contrast to the NSGA algorithm, NSGA-II additionally diminishes computational complexity.

This study tackles the multi-objective challenge of optimizing power generation and shipping for the Shatuo Hydropower Station through the integration of the GA-NSGA-II algorithm. Initially, we encode the downstream flow of the Shatuo Hydropower Station for each time period and generate scheduling outcomes using the GA algorithm. These results from the GA algorithm serve as preliminary insights, and the subsequent application of the NSGA-II algorithm refines the solution, ultimately achieving a scheduling outcome that aligns with both generation and shipping objectives. The algorithm leverages the diversity benefits inherent in NSGA-II individuals, mitigating premature convergence resulting from excessive genetic inheritance. The incorporation of the GA algorithm contributes to improving the precision of the optimization results obtained through the NSGA-II algorithm. The detailed flowchart outlining the algorithm is depicted in Fig. 7. The sequential steps of the algorithm include:

Step 1: Utilize historical load data as input and conduct experimental analysis employing various forecasting algorithms to determine the algorithm exhibiting the highest forecasting accuracy. Subsequently, proceed to forecast power load. While, in the current practical case, the CNN-GRU model was chosen for the prediction task, the proposed GA-NSGA-II algorithm can be generalized for different datasets by selecting the most suitable model from a few available options.
Step 2: Examine the constraints inherent in the practical scheduling procedure of the hydropower station.
Step 3: Set the initial downstream flow values for each time period, taking into account the constraints found in Step 2.
Step 4: Apply the GA algorithm to produce a set of enhanced solutions. Then, use the optimal solutions derived from GA as prior knowledge for NSGA-II and initiate the algorithm.
Step 5: Verify if the initial generation of scheduling outcomes has been produced. If not, create the first generation of subgroups through non-dominated sorting, selection, crossover, and mutation.
Step 6: Verify the successful merging of parent and child populations. If unsuccessful, conduct operations like fast non-dominated sorting, calculating crowding, and employing the elite strategy to generate new parent populations. Subsequently, carry out selection, crossover, and mutation on the newly generated populations.
Step 7: Check if the number of iterations fulfills the specified condition. If not, continue iterating through the loop until the condition is satisfied. This iterative process produces the scheduling scheme with the optimal discharge flow value for each time period of the station.

## 5. Experiments

The experiments are divided into two segments: power load forecasting and optimal scheduling. To predict the power loads, the seven models mentioned earlier are utilized to forecast power loads, and the most accurate prediction algorithm is determined through result comparisons. In the experiments for optimal scheduling, we employ both the predicted and the actual power loads as input. The performance of the GA-NSGA-II algorithm is compared to that of the NSGA-II algorithm, and the distinctions between the two approaches are examined. Additionally, experiments are conducted using actual load data with the NSGA-II algorithm, and a comparison is performed to identify differences between the NSGA-II and GA-NSGA-II algorithms.

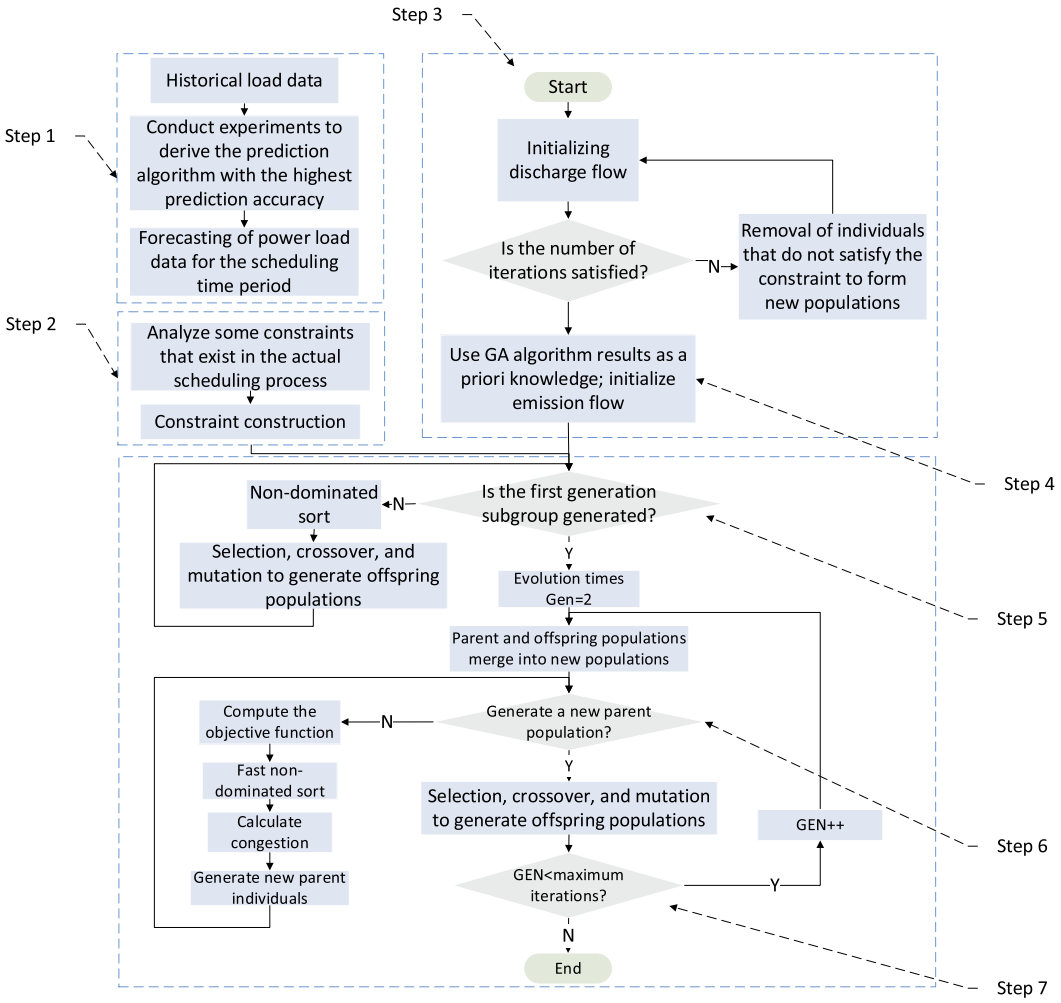



**Fig. 7.** The GA-NSGA-II algorithm.

**Table 2**
Parameters of the experimental environment.

| Hardware and software environment | Configuration |
|---|---|
| CPU | 12th Gen Intel(R) Core(TM) i7-12700H 2.70 GHz |
| RAM | DDR5 16G 4800MHZ |
| Hard Disk | SAMSUNG MZVL2512HCJQ-00BL2 |
| OS | Windows11 |
| Programming Environment | Python3.9 |
| Development language | Python |

### 5.1. Setups

The experimental hardware and software environments are outlined in Table 2.

### 5.2. Testing data

In the practical case study, a comprehensive dataset comprising 51,840 data points was collected from October to November 2021 for the Shatuo Hydropower Station. This dataset includes information on tailwater level, flow, load, and upstream water level. The dataset was employed in performing the power load prediction experiment, as specified in Table 3.

**Table 3**
Shatuo Hydropower Station sample data.

| Tailwater level (m) | Discharge flow ($m^3/s$) | Load (kW) | Upstream water level (m) |
|---|---|---|---|
| 289.81 | 311.82 | 202.31 | 361.02 |
| 289.82 | 313.19 | 204.81 | 361.02 |
| 289.76 | 311.04 | 206.00 | 361.02 |
| 289.78 | 316.88 | 203.20 | 361.02 |
| 289.79 | 313.97 | 203.31 | 361.02 |
| 289.84 | 313.75 | 205.60 | 361.02 |
| ... | ... | ... | ... |
| 289.78 | 312.19 | 204.70 | 361.02 |

**Table 4**
Scheduling data.

| Load (kW) | Flow ($m^3/s$) | Flow_down ($m^3/s$) | Flow_up ($m^3/s$) | G | H (m) |
|---|---|---|---|---|---|
| 199.2 | 332 | 327 | 361.02 | 0.96 | 0.065687 |
| 201.5 | 334.8 | 329.8 | 361.02 | 0.96 | 0.065752 |
| 357.5 | 619.1 | 614.1 | 361.02 | 0.84 | 0.062974 |
| 388.4 | 665.3 | 661.3 | 361.02 | 0.82 | 0.062413 |
| 410.8 | 702.3 | 697.3 | 361.02 | 0.80 | 0.559213 |
| 205.8 | 337 | 332 | 361.02 | 0.96 | 0.055074 |
| ... | ... | ... | ... | ... | ... |
| 250.5 | 336.7 | 331.7 | 361.02 | 0.96 | 0.055456 |

The sampling interval for data points is one minute; nevertheless, scheduling every minute during the scheduling process is not feasible. Consequently, the data is smoothed, and an average value is computed per quarter-hour (15 minutes). Scheduling experiments are then performed based on 96 data points over a 24-hour day. The data is classified into power load (Load), real-time flow ($Flow$), upper flow limit ($Flow_up$), lower flow limit ($Flow_down$), navigation rate ($G$), and hydraulic head ($H$), as depicted in Table 4.

### 5.3. The findings for power load prediction

In the SVR model experiments, the data is processed as follows. Since the collected data only contains timestamps and load values, creating a time series dataset without feature data, we shift the load data down by 33% to generate a new column labeled as $PLoad$. Subsequently, the dataset is divided into a 33.3% test set and a 66.6% training set. For the experiments, we choose linear, Gaussian, Sigmoid, and polynomial kernel functions for prediction. Ten experiments are conducted for each kernel function to compute the average of each metric. Finally, we visualize the top 1,200 true and predicted values of the test set for output. We employ evaluation metrics such as Mean Squared Error Loss (MSE), Mean Absolute Error Loss (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-Squared (R2) to assess the experimental results. The specific evaluation metrics are detailed in Table 5, illustrating the diverse performance aspects of the algorithms. MSE quantifies the discrepancy between the observed and true values. MAE calculates the average of the absolute error between the observed and true values. RMSE is the square root of the ratio of the squared deviation of the predicted value from the true value to the number of observations. MAPE is a comparison with MAE, introducing an additional true value in the denominator below the difference between the predicted and true values. For R2, the numerator part signifies the sum of the squared differences between the true and predicted values, while the denominator part denotes the sum of the squared differences between the true and mean values. The experimental outcomes are illustrated in Figs. 8 to 11 and Table 6, with green lines representing the actual power loads and red lines representing the forecasted power loads.

The ARIMA model partitions the dataset into training and test sets as follows: The training set comprises data from 2021-10-01 00:00 to 2021-10-28 00:00, while the test set encompasses data from 2021-10-29 00:00 to 2021-11-05 23:59. The order parameters in ARIMA are determined by analyzing trailing Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots, resulting in the selection of values $P$, $D$, and $Q$ as 2, 1, and 2, respectively.

Following that, we present a visualization of the initial 1,200 real and predicted values from the test set, where the green line denotes the actual power loads, and the red line denotes the forecasted power loads. This is depicted in Fig. 12. The predicted data closely match the real data, suggesting a strong fitting effect.

The CNN model divides the training set, allocating one-third of the total data with a test size of 0.33. Before training, the data undergoes normalization. The model architecture consists of two convolutional layers linked to a pooling layer, succeeded by a fully connected layer. In the initial convolutional kernel, the number of filters is configured as 8, the time domain window length is set to 3 (kernel_size), and the activation function is defined as “relu”. Likewise, for the second convolutional kernel, the number of filters

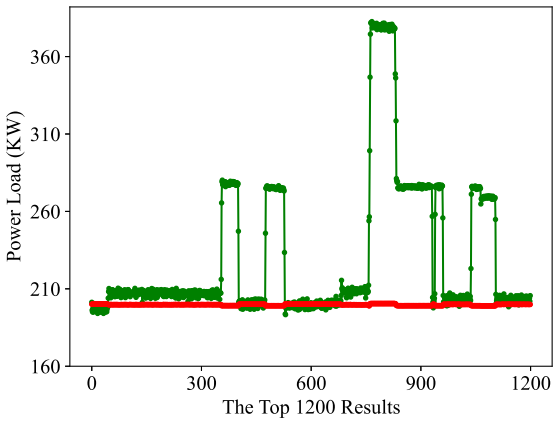



**Fig. 8.** Linear kernel function.

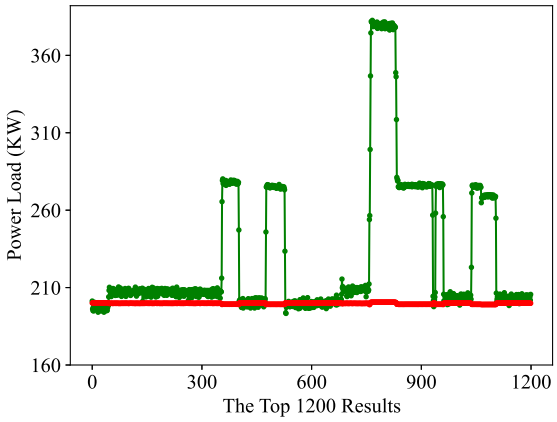



**Fig. 9.** Gaussian kernel function.

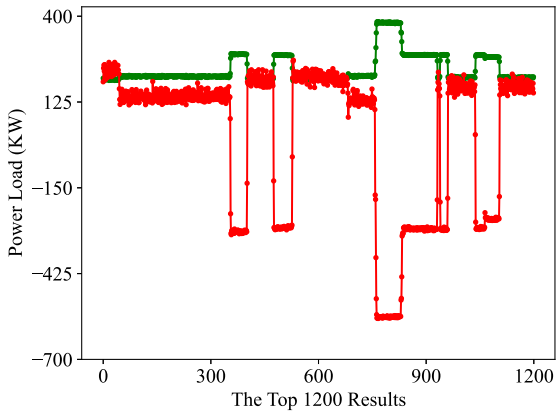



**Fig. 10.** Sigmod kernel function.

is set to 16, the time domain window length is 3 (kernel_size), and the activation function is specified as “relu”. The pooling layer incorporates a pooling window size (pool_size) of 2. The training procedure consists of 50 iterations with a batch size gradient of 100. Ultimately, the initial 1,200 values from the test set are subjected to inverse normalization and presented visually. The actual power load is depicted by the blue line, while the red line signifies the predicted power load. The experimental outcomes are depicted in Fig. 13. The predicted data exhibit a close resemblance to the actual data, suggesting a well-fitting effect.

The LSTM model divides the training set, using one-third of the total data with a test_size of 0.33. Initially, two connected LSTM layers are followed by a dimensionality reduction operation returning a one-dimensional array. Subsequently, two additional LSTM layers with 32 units each are connected, along with four feedforward network layers featuring 32 neurons each. The selected loss

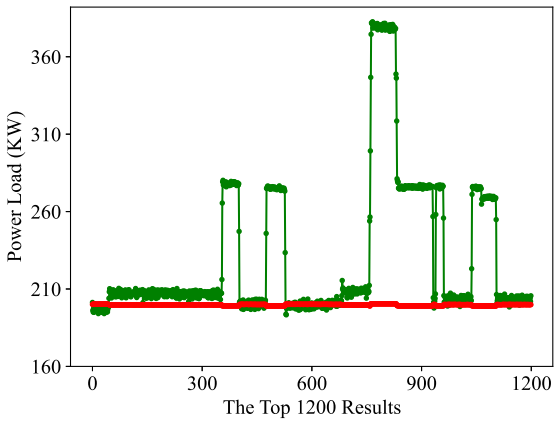



**Fig. 11.** Polynomial kernel function.

**Table 5**
Evaluation Metrics for Predictive Algorithms.

| Performance Metrics | Optimal value | Mathematical formula |
|---|---|---|
| MSE: Assesses the extent of data variability, with a smaller value indicating higher model accuracy. | 0 | $MSE = \frac{1}{n}\sum_{i=1}^{T}\left(\hat{y}_i - y_i\right)^2$ |
| MAE: The mean distance between the predicted value of the model and the true value in the sample. | 0 | $MAE = \frac{1}{n}\sum_{i=1}^{n}\left\|\hat{y}_i - y_i\right\|$ |
| RMSE: Assesses the difference between observed and actual values, with smaller values indicating improved prediction accuracy. | 0 | $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}$ |
| MAPE: A statistical indicator employed to gauge prediction accuracy, where 0% signifies a flawless model, and values exceeding 100% indicate a suboptimal model. | 0 | $MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left\|\frac{\hat{y}_i - y_i}{y_i}\right\|$ |
| R2: A metric indicating the degree of model fit, ranging between 0 and 1. Higher values indicate a better fit of the model. | 1 | $R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$ |

**Table 6**
SVR experimental results comparison.

| SVR model kernel function | MSE | MAE | RMSE | MAPE | R2 |
|---|---|---|---|---|---|
| Linear kernel | 3111.8 | 19.8 | 55.7 | 9.8 | -0.12278 |
| Gaussian kernel | 3113.9 | 19.8 | 55.8 | 9.8 | -0.12353 |
| Sigmod kernel | 75072.5 | 130.9 | 273.9 | 85.7 | -26.0868 |
| polynomial kernel | 3085.7 | 19.6 | 55.5 | 9.8 | -0.11336 |

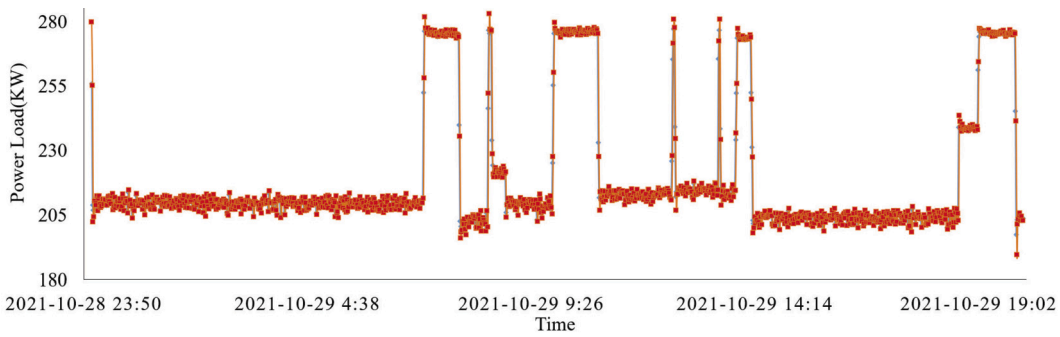



**Fig. 12.** ARIMA model prediction.

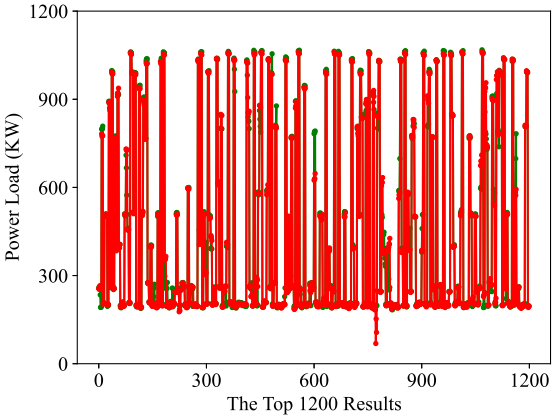



**Fig. 13.** CNN model prediction.

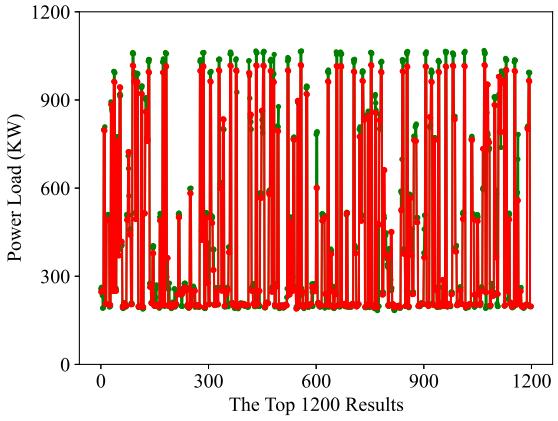



**Fig. 14.** LSTM model prediction.

function is the mean square error. Finally, the initial 1,200 real and predicted values of the test set are visualized for output. The green line illustrates the actual power load, while the red line represents the predicted power load. The predicted outcomes are portrayed in Fig. 14, revealing a close alignment between the predicted and real data, indicating a well-fitting effect.

The GRU model shares structural similarities with the LSTM model, incorporating the same parameter selection. It comprises two GRU layers, succeeded by a dimensionality reduction operation returning a one-dimensional array. Subsequently, the model establishes connections with two additional GRU layers. The employed loss function is also the mean squared error. Finally, the initial 1,200 real and predicted values of the test set are visualized. The green line depicts the actual power load, while the red line represents the predicted power load. The predicted outcomes are depicted in Fig. 15, showcasing a close resemblance between the predicted and real data and underscoring a commendable fitting effect.

The integration of CNN with LSTM and GRU models is employed for power load forecasting, with each model retaining identical parameter settings. Subsequently, the initial 1,200 real and predicted values of the test set are visualized for output. The green line signifies the actual power load, while the red line represents the predicted power load. The prediction results are illustrated in Figs. 16 and 17, showcasing a close alignment between the predicted and real data and underscoring a robust fitting effect. The specific values of evaluation metrics are outlined in Table 8.

Throughout our experiments, the GRU model demonstrates comparable performance to the CNN-GRU model. Despite this, the CNN-GRU model proves to be more adept at processing sequence data, enabling it to extract more intricate feature information while retaining the nuances of the sequence. Consequently, we opt for the CNN-GRU model in predicting power load data, yielding the associated results.

### *5.4. The findings for the scheduling algorithms*

In our experimentation, we perform separate scheduling trials using the NSGA-II algorithm for both predicted and actual load data. The primary aim of scheduling is to minimize the objective function $F1$ while maximizing the value of objective function $F2$. We meticulously compare the outcomes of both sets of experiments and incorporate experiments utilizing the NSGA-II algorithm as a point of reference.

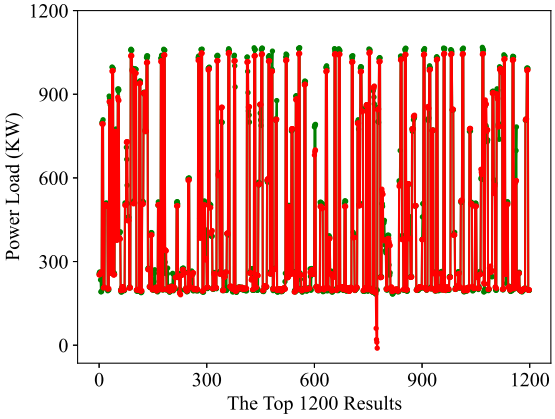



**Fig. 15.** GRU model prediction.

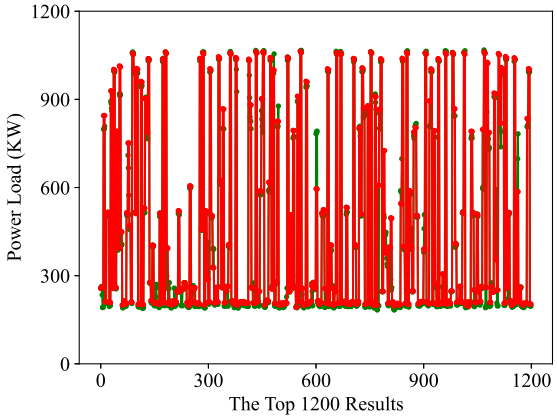



**Fig. 16.** CNN-LSTM model prediction.

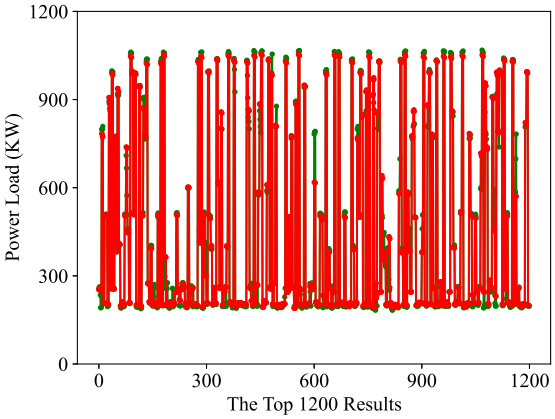



**Fig. 17.** CNN-GRU model prediction.

We choose the dataset from the Shatuo Hydropower Station on November 1 for our experimental analysis. Initially, scheduling occurs every fifteen minutes, and the simulated scheduling scheme is executed every six hours. The experiments utilize the GA-NSGA-II algorithm. In the first layer of the GA algorithm, the parameter NIMD is configured as 800, the maximum evolutionary generation MAXGEN is set to 20, and the population size is 16,000. In the second layer of the NSGA-II algorithm, the parameter NIMD is adjusted to 30, the maximum evolutionary generation MAXGEN is defined as 50, and the population size is set at 1,500. Since the scheduling data involves real values, the encoding method Encoding is designated as RI.

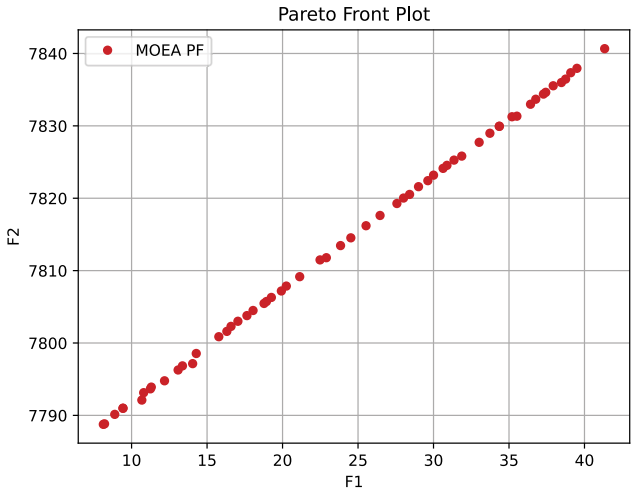



**Fig. 18.** Illustrates the scheduling results from 0:00 to 6:00.

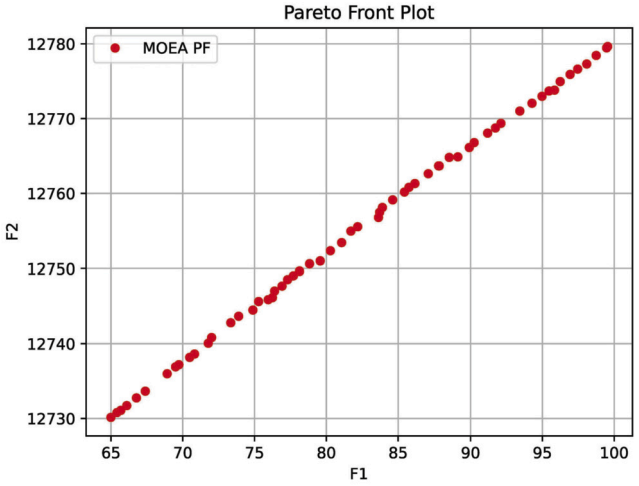



**Fig. 19.** Illustrates the scheduling results from 6:00 to 12:00.

The experimental outcomes are graphically represented, with each red dot on the plot corresponding to a population (feasible solution) – denoting the discharge flow value meeting the condition for each time period. The coordinates denote the values of the objective functions $F1$ and $F2$ for each population, respectively. As the scheduling objective aims to optimize $F1$ for efficient generation and $F2$ for efficient navigation, individuals closer to the top-left corner indicate better results and scheduling outcomes. The results for prediction scheduling are showcased in Figs. 18-21, while those for actual load scheduling are presented in Figs. 22-25. Additionally, the results using the NSGA-II algorithm are depicted in Figs. 26-29. A solution displaying high adaptability from the GA-NSGA-II model is specifically chosen to illustrate the discharge flow data, as outlined in Table 7.

### 5.5. Discussion of the experimental results

Following the training of the model with historical data, the test data is introduced to the trained model to compute forecasted values. These predicted values for the test data are subsequently juxtaposed with their corresponding true values to assess model performance. This study utilizes five regression metrics for evaluation. Regarding these metrics, $n$ denotes the number of samples, $y_i$ represents the true value of the $i$-th sample, $\hat{y}_i$ indicates the predicted value of the $i$-th sample, and $\overline{y}$ signifies the average of the true values. The values for the evaluation metrics are presented in Table 8.

Due to the absence of essential feature attributes in our dataset for SVR, a novel label is created through the shift function, resulting in suboptimal performance of the SVM model. Conversely, the ARIMA model surpasses SVM notably, achieving an R2 metric of 0.95. The GRU model and the integrated CNN-GRU model emerge as top performers, each achieving an R2 metric of 0.99 or higher, as indicated in Table 8.

With a prediction accuracy of around 99% using the CNN-GRU model, the discrepancy between our scheduling based on predicted power load data and the actual scheduling scenario is correspondingly minimal. As shown in Table 5, the difference in downstream

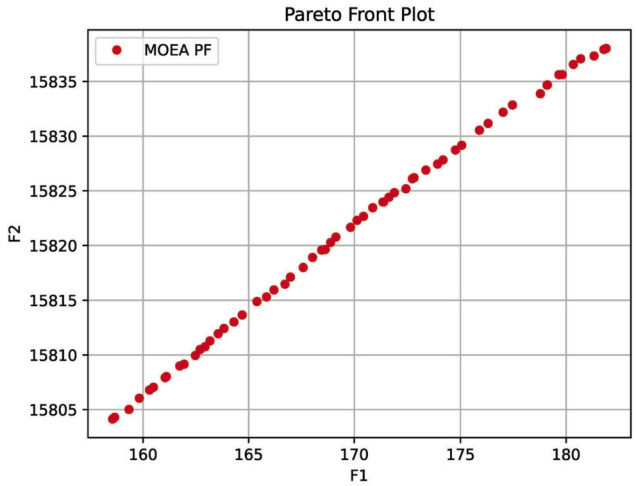



**Fig. 20.** Illustrates the scheduling results from 12:00 to 18:00.

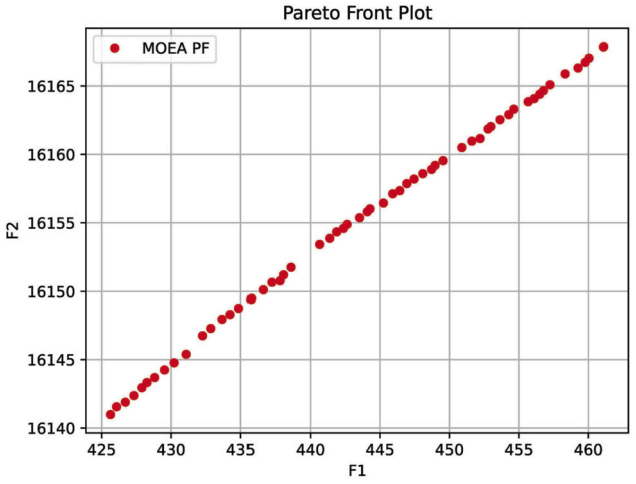



**Fig. 21.** Illustrates the scheduling results from 18:00 to 24:00.

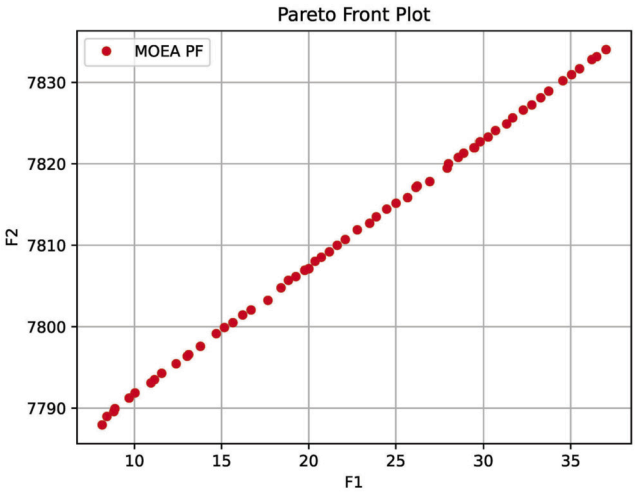



**Fig. 22.** Illustrates the scheduling results from 0:00 to 6:00.

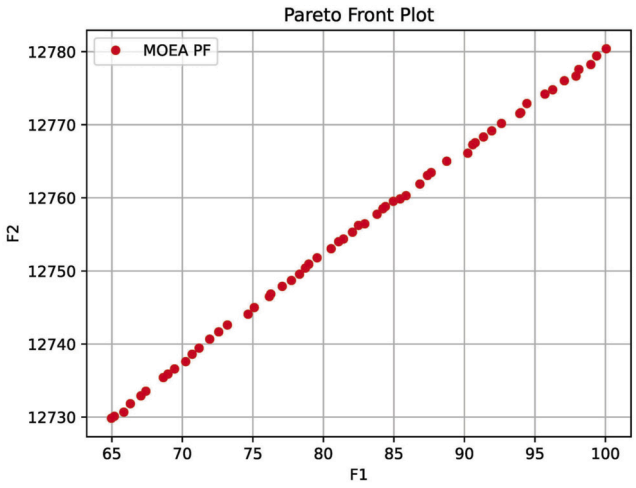



**Fig. 23.** Illustrates the scheduling results from 6:00 to 12:00.

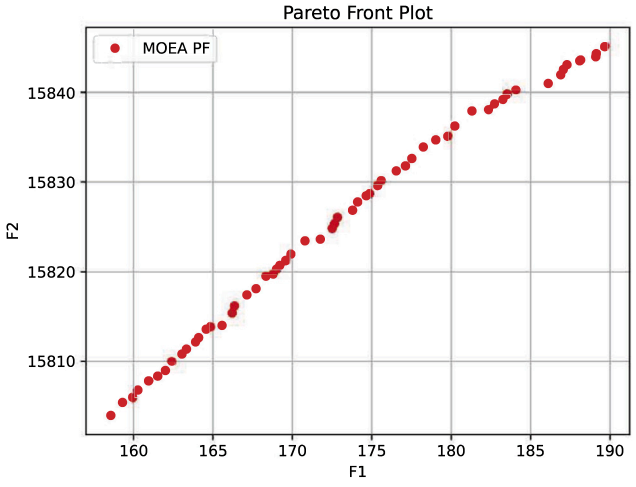



**Fig. 24.** Illustrates the scheduling results from 12:00 to 18:00.

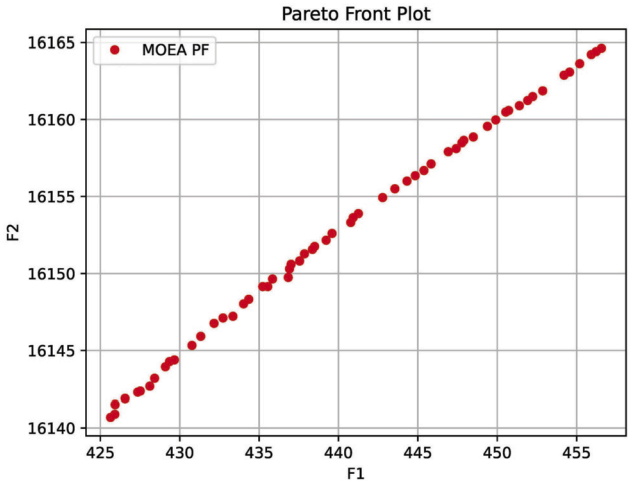



**Fig. 25.** Illustrates the scheduling results from 18:00 to 24:00.

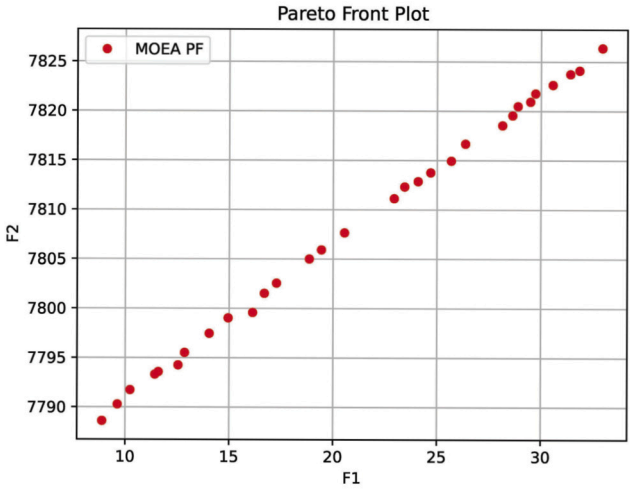



**Fig. 26.** Illustrates the scheduling results from 0:00 to 6:00.

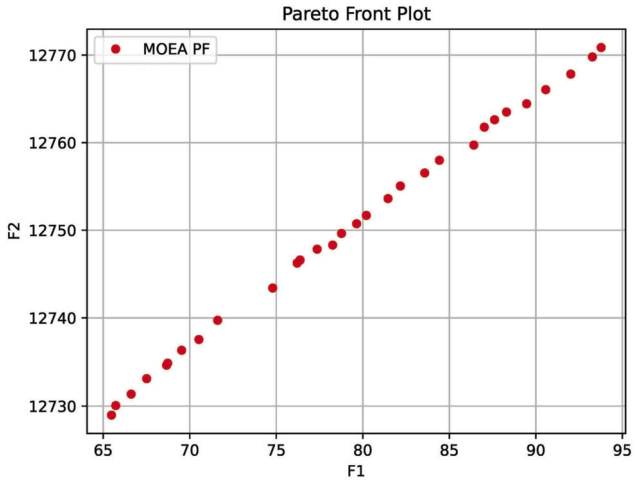



**Fig. 27.** Illustrates the scheduling results from 6:00 to 12:00.

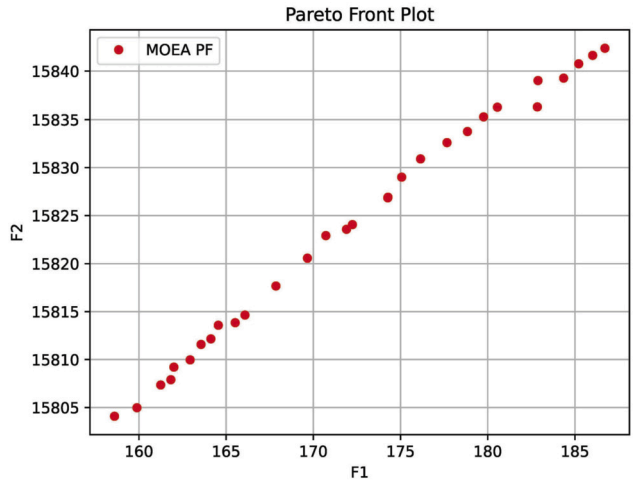



**Fig. 28.** Illustrates the scheduling results from 12:00 to 18:00.

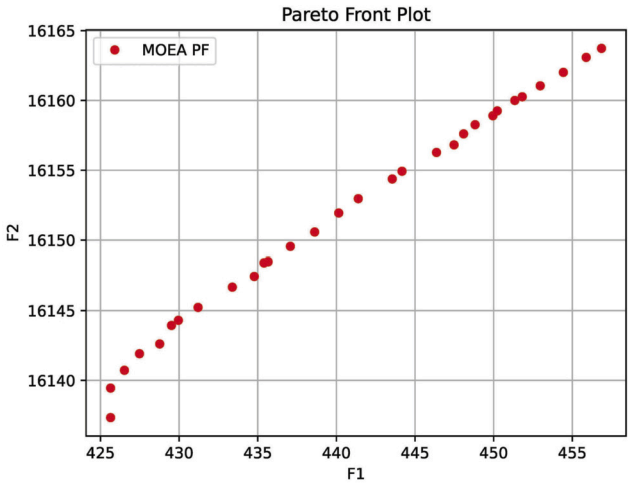



**Fig. 29.** Illustrates the scheduling results from 18:00 to 24:00.

**Table 7**
Discharge flow data for scheduling experiments at Shatuo Hydropower Station.

| Time Period | discharge flow | | Time Period | discharge flow | | Time Period | discharge flow | | Time Period | discharge flow | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Predicted load | Actual load | | Predicted load | Actual load | | Predicted load | Actual load | | Predicted load | Actual load |
| 00:15 | 327 | 327 | 06:15 | 348 | 348 | 12:15 | 333 | 332 | 18:15 | 1503 | 1503 |
| 00:30 | 335 | 335 | 06:30 | 615 | 614 | 12:30 | 331 | 331 | 18:30 | 1498 | 1498 |
| 00:45 | 329 | 330 | 06:45 | 660 | 660 | 12:45 | 508 | 508 | 18:45 | 1403 | 1402 |
| 01:00 | 329 | 333 | 07:00 | 542 | 541 | 13:00 | 699 | 699 | 19:00 | 1638 | 1637 |
| 01:15 | 329 | 332 | 07:15 | 448 | 447 | 13:15 | 1365 | 1365 | 19:15 | 1667 | 1669 |
| 01:30 | 330 | 330 | 07:30 | 451 | 450 | 13:30 | 1414 | 1419 | 19:30 | 1672 | 1671 |
| 01:45 | 338 | 339 | 07:45 | 460 | 461 | 13:45 | 1106 | 1104 | 19:45 | 1821 | 1821 |
| 02:00 | 339 | 340 | 08:00 | 458 | 458 | 14:00 | 1160 | 1160 | 20:00 | 1876 | 1877 |
| 02:15 | 340 | 341 | 08:15 | 458 | 460 | 14:15 | 1232 | 1232 | 20:15 | 1869 | 1872 |
| 02:30 | 339 | 342 | 08:30 | 468 | 469 | 14:30 | 1378 | 1377 | 20:30 | 1872 | 1873 |
| 02:45 | 342 | 342 | 08:45 | 641 | 644 | 14:45 | 1370 | 1370 | 20:45 | 1871 | 1872 |
| 03:00 | 340 | 341 | 09:00 | 707 | 707 | 15:00 | 1432 | 1434 | 21:00 | 1871 | 1871 |
| 03:15 | 341 | 341 | 09:15 | 1271 | 1272 | 15:15 | 1606 | 1608 | 21:15 | 1876 | 1876 |
| 03:30 | 341 | 342 | 09:30 | 1143 | 1144 | 15:30 | 1869 | 1870 | 21:30 | 1552 | 1552 |
| 03:45 | 339 | 342 | 09:45 | 1197 | 1197 | 15:45 | 1860 | 1860 | 21:45 | 1381 | 1382 |
| 04:00 | 342 | 342 | 10:00 | 1389 | 1392 | 16:00 | 1866 | 1869 | 22:00 | 1580 | 1580 |
| 04:15 | 338 | 342 | 10:15 | 1195 | 1195 | 16:15 | 1741 | 1744 | 22:15 | 1879 | 1881 |
| 04:30 | 339 | 341 | 10:30 | 1006 | 1006 | 16:30 | 1882 | 1882 | 22:30 | 1682 | 1682 |
| 04:45 | 343 | 343 | 10:45 | 1008 | 1009 | 16:45 | 1878 | 1878 | 22:45 | 1604 | 1604 |
| 05:00 | 342 | 342 | 11:00 | 979 | 979 | 17:00 | 1619 | 1623 | 23:00 | 1397 | 1398 |
| 05:15 | 341 | 341 | 11:15 | 662 | 664 | 17:15 | 1441 | 1443 | 23:15 | 1109 | 1110 |
| 05:30 | 339 | 341 | 11:30 | 357 | 557 | 17:30 | 1419 | 1418 | 23:30 | 1108 | 1109 |
| 05:45 | 340 | 341 | 11:45 | 338 | 338 | 17:45 | 1388 | 1387 | 23:45 | 690 | 690 |
| 06:00 | 339 | 340 | 12:00 | 339 | 340 | 18:00 | 1393 | 1282 | 24:00 | 341 | 341 |

The unit of discharge flow in the table is $m^3/s$. The unit of the predicted load and actual load is kW.

flows derived from the two experiments is so negligible that it has virtually no impact on the real scheduling process. The difference between the two stays within 5 for each of the 96 scheduling instances in a day. The overall variation in the objective function values is also within single digits. Scheduling simulations are conducted every six hours, and the objective function values for each six-hour simulation are outlined in Table 9.

We performed experiments to assess the effectiveness of the GA-NSGA-II algorithm in optimizing dual objectives with multiple constraints related to power generation and shipping for hydropower stations, comparing its performance with the NSGA-II algorithm.

**Table 8**
Comparison of prediction models.

| Performance Metrics | MSE | MAE | RMSE | MAPE | R2 |
|---|---|---|---|---|---|
| SVR(poly) | 3082.969 | 19.651 | 55.524 | 9.820 | 0.1123 |
| ARIMA | 37.158 | 2.629 | 6.0957 | 24.432 | 0.950 |
| CNN | 0.001182 | 0.0125 | 0.0344 | 3.138 | 0.985 |
| GRU | 0.000583 | 0.01021 | 0.0241 | 2.9532 | 0.993 |
| LSTM | 0.001095 | 0.0146 | 0.0331 | 3.8749 | 0.986 |
| CNN-GRU | 0.0006796 | 0.0116477 | 0.0260 | 3.1062 | 0.991 |
| CNN-LSTM | 0.0008309 | 0.0132209 | 0.0288 | 3.6110 | 0.9895 |

**Table 9**
Scheduling objective function values.

| Time | Forecast load scheduling objective function value | | Actual load scheduling objective function value | |
|---|---|---|---|---|
| | Objective function F1 | Objective function F2 | Objective function F1 | Objective function F2 |
| 00:00-06:00 | 14.19 | 7778.80 | 11.17 | 7793.83 |
| 06:00-12:00 | 66.79 | 12720.40 | 64.96 | 12726.78 |
| 12:00-18:00 | 160.49 | 15796.74 | 159.04 | 15800.03 |
| 18:00-24:00 | 428.74 | 16136.95 | 426.74 | 16121.11 |

**Table 10**
Scheduling model comparison.

| Time | HV index of GA-NSGA-II | HV index of NSGA-II | Spacing index of GA-NSGA-II | Spacing index of NSGA-II | Running time of GA-NSGA-II (t/s) | Running time of NSGA-II (t/s) |
|---|---|---|---|---|---|---|
| 00:00-06:00 | 0.076 | 0.067 | 0.37 | 0.40 | 0.040 s | 0.13 s |
| 06:00-12:00 | 0.035 | 0.031 | 0.34 | 0.45 | 0.030 s | 0.13 s |
| 12:00-18:00 | 0.021 | 0.020 | 0.41 | 0.43 | 0.030 s | 0.13 s |
| 18:00-24:00 | 0.015 | 0.013 | 0.32 | 0.45 | 0.027 s | 0.13 s |

Utilizing actual load data, the objective was to minimize the value of the objective function $F1$ and maximize the value of the objective function $F2$. The experimental outcomes emphasize that the GA-NSGA-II algorithm is better suited for our scheduling requirements, as it identifies a solution set that closely aligns with the upper-left region of the coordinate axis. We also evaluated the effectiveness of the two algorithms using the Hypervolume Metric ($HV$) and Spacing. $HV$ gauges the disparity between the solution set and the true value, providing insights into the convergence of the solution set. Spacing evaluates the internal distribution of the solution set, emphasizing its uniformity. Enhanced $HV$ values signify improved convergence of the solution set, whereas reduced Spacing values suggest a more uniform distribution. The outcomes are presented in Table 10, showcasing the GA-NSGA-II algorithm's superior convergence and uniformity in comparison to the NSGA-II algorithm.

In summary, our experimental findings suggest that the GA-NSGA-II algorithm proves to be a superior option for optimizing the dual goals of hydrapower generation and ship navigation within the practical hydropower station. Furthermore, its versatility extends to other scheduling challenges featuring similar objectives.

## 6. Conclusion

Addressing the challenge of delayed power load data collection at Shatuo Hydropower Station, we introduce a joint scheduling model for the hydropower station incorporating power load prediction. Following a comparative analysis, the CNN-GRU prediction algorithm is chosen for practical implementation. The outcomes reveal a fitting accuracy of approximately 0.99. The variation between scheduling using predicted load data and actual results is limited to within 5 $m^3/s$, aligning with the practical scheduling needs. This investigation employed the GA-NSGA-II algorithm to enhance the scheduling of the Shatuo Hydropower Station, targeting the dual objectives of hydrapower generation and ship navigation. In the four scheduling time periods, the objective function values, HV, spacing, and running time surpassed those achieved by the NSGA-II algorithm. Moreover, the prediction-based scheduling model exhibits versatile applicability and is not confined to the Shatuo Hydropower Station. With minor adjustments, it can be adapted for other stations, offering a novel approach to address multi-objective scheduling challenges when timely data availability is a constraint.

There are still limitations to our study. We have only examined the dual benefits of hydrapower generation and ship navigation in the scheduling optimization of the Shatuo Hydropower Station. Moving forward, we plan to take into account additional factors,

such as the ecological impact of the Shatuo Hydropower Station's operation. We aim to further enhance and optimize the scheduling results by incorporating new algorithms.

## CRediT authorship contribution statement

**Guangqin Huang:** Project administration, Methodology, Investigation, Data curation. **Ming Tan:** Supervision, Resources, Conceptualization. **Zhihang Meng:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis. **Jiaqi Yan:** Validation, Resources. **Jin Chen:** Supervision, Resources, Funding acquisition, Conceptualization. **Qiang Qu:** Writing – review & editing, Supervision, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The data that support the findings of this study are available on request from the corresponding author.

## Acknowledgements

This work was supported by the Science and Technology Project of the Department of Transportation Science and Technology of Guizhou Province (No.2022-221-008) and the Shenzhen Basic Research Special Basic Research Key Project (JCYJ20200109115422828).

## References

[1] T. Ackermann, D.P. Loucks, D. Schwanenberg, M. Detering, Real-time modeling for navigation and hydropower in the River Mosel, J. Water Resour. Plan. Manag. 126 (5) (2000) 298–303, https://doi.org/10.1061/(ASCE)0733-9496(2000)126:5(298).
[2] E.B. Tirkolaee, A. Goli, M. Hematian, A.K. Sangaiah, T. Han, Multi-objective multi-mode resource constrained project scheduling problem using Pareto-based algorithms, Computing 101 (2019) 547–570, https://doi.org/10.1007/s00607-018-00693-1.
[3] D. Alberg, M. Last, Short-term load forecasting in smart meters with sliding window-based ARIMA algorithms, Vietnam J. Comput. Sci. 5 (2018) 241–249, https://doi.org/10.1007/s40595-018-0119-7.
[4] Y. Chen, P. Xu, Y. Chu, W. Li, Y. Wu, L. Ni, Y. Bao, K. Wang, Short-term electrical load forecasting using the support vector regression (SVR) model to calculate the demand response baseline for office buildings, Appl. Energy 195 (2017) 659–670, https://doi.org/10.1016/j.apenergy.2017.03.034.
[5] L. Xu, C. Li, X. Xie, G. Zhang, Long-short-term memory network based hybrid model for short-term electrical load forecasting, Information 9 (7) (2018) 165, https://doi.org/10.1016/j.apenergy.2017.03.034.
[6] B. Liu, C. Fu, A. Bielefield, Y.Q. Liu, Forecasting of Chinese primary energy consumption in 2021 with GRU artificial neural network, Energies 10 (10) (2017) 1453, https://doi.org/10.3390/en10101453.
[7] A.M. Tudose, D.O. Sidea, I.I. Picioroaga, V.A. Boicea, C. Bulac, A CNN based model for short-term load forecasting: a real case study on the Romanian power system, in: 2020 55th International Universities Power Engineering Conference (UPEC), IEEE, 2020, pp. 1–6.
[8] Z. Cheng, L. Wang, Y. Yang, A hybrid feature pyramid CNN-LSTM model with seasonal inflection month correction for medium- and long-term power load forecasting, Energies 16 (7) (2023) 3081, https://doi.org/10.3390/en16073081.
[9] M. Sajjad, Z.A. Khan, A. Ullah, T. Hussain, W. Ullah, M.Y. Lee, S.W. Baik, A novel CNN-GRU-based hybrid approach for short-term residential load forecasting, IEEE Access 8 (2020) 143759–143768, https://doi.org/10.3390/en16073081.
[10] S. Katoch, S.S. Chauhan, V. Kumar, A review on genetic algorithm: past, present, and future, Multimed. Tools Appl. 80 (2021) 8091–8126, https://doi.org/10.1007/s11042-020-10139-6.
[11] S. Verma, M. Pant, V. Snasel, A comprehensive review on NSGA-II for multi-objective combinatorial optimization problems, IEEE Access 9 (2021) 57757–57791, https://doi.org/10.1109/ACCESS.2021.3070634.
[12] I. Rahimi, A.H. Gandomi, K. Deb, F. Chen, M.R. Nikoo, Scheduling by NSGA-II: review and bibliometric analysis, Processes 10 (1) (2022) 98, https://doi.org/10.3390/pr10010098.
[13] H. Borhanazad, S. Mekhilef, V.G. Ganapathy, M. Modiri-Delshad, A. Mirtaheri, Optimization of micro-grid system using MOPSO, Renew. Energy 71 (2014) 295–306, https://doi.org/10.1016/j.renene.2014.05.006.
[14] T. Murata, H. Ishibuchi, et al., MOGA: Multi-Objective Genetic Algorithms, IEEE International Conference on Evolutionary Computation, vol. 1, IEEE, Piscataway, 1995, pp. 289–294.
[15] W. Liu, F. Zhu, J. Chen, H. Wang, B. Xu, P. Song, P.-a. Zhong, X. Lei, C. Wang, M. Yan, et al., Multi-objective optimization scheduling of wind–photovoltaic–hydropower systems considering riverine ecosystem, Energy Convers. Manag. 196 (2019) 32–43, https://doi.org/10.1016/j.enconman.2019.05.104.
[16] X. Meng, J. Chang, X. Wang, Y. Wang, Multi-objective hydropower station operation using an improved cuckoo search algorithm, Energy 168 (2019) 425–439, https://doi.org/10.1016/j.energy.2018.11.096.
[17] Z. Yang, K. Yang, L. Su, H. Hu, The multi-objective operation for cascade reservoirs using MMOSFLA with emphasis on power generation and ecological benefit, J. Hydroinform. 21 (2) (2019) 257–278, https://doi.org/10.2166/hydro.2019.064.
[18] R. Fang, Z. Popole, Multi-objective optimized scheduling model for hydropower reservoir based on improved particle swarm optimization algorithm, Environ. Sci. Pollut. Res. 27 (2020) 12842–12850, https://doi.org/10.1007/s11356-019-04434-5.
[19] Q. Shen, L. Mo, G. Liu, Hybrid optimize method for multi-objective scheduling of hydropower stations with coupled power generation, navigation demand and ecological benefit, J. Coast. Res. 104 (SI) (2020) 379–384, https://doi.org/10.2112/JCR-SI104-067.1.
[20] Y. Feng, J. Xu, Y. Hong, Y. Wang, Z. Yuan, C. Wang, Reservoir scheduling using a multi-objective cuckoo search algorithm under climate change in Jinsha River, China, Water 13 (13) (2021) 1803, https://doi.org/10.3390/w13131803.
[21] F. Kong, Y. Zhuo, C. Song, Ecological multi-objective joint optimized scheduling of cascade hydropower plants based on improved marine predators algorithm, J. Renew. Sustain. Energy 14 (4) (2022) 044503, https://doi.org/10.1063/5.0091309.

[22] Y. Chen, M. Wang, Y. Zhang, Y. Lu, B. Xu, L. Yu, Cascade Hydropower System Operation Considering Ecological Flow Based on Different Multi-Objective Genetic Algorithms, Water Resources Management, 2023, pp. 1–18.
[23] E. Zhou, X. Liu, Z. Meng, S. Yu, J. Mei, Q. Qu, Hydropower station scheduling with ship arrival prediction and energy storage, Sci. Rep. 13 (2023) 18969, https://doi.org/10.1038/s41598-023-45995-3.
[24] A.J. Smola, B. Schölkopf, A tutorial on support vector regression, Stat. Comput. 14 (2004) 199–222, https://doi.org/10.1023/B:STCO.0000035301.49549.88.
[25] R.H. Shumway, D.S. Stoffer, R.H. Shumway, D.S. Stoffer, ARIMA models, Time Ser. Anal. Appl. R Exampl. (2017) 75–163, https://doi.org/10.1007/978-3-319-52452-8_3.
[26] L. Alzubaidi, J. Zhang, A.J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M.A. Fadhel, M. Al-Amidie, L. Farhan, Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, J. Big Data 8 (2021) 1–74, https://doi.org/10.1186/s40537-021-00444-8.
[27] Y. Yu, X. Si, C. Hu, J. Zhang, A review of recurrent neural networks: LSTM cells and network architectures, Neural Comput. 31 (7) (2019) 1235–1270, https://doi.org/10.1162/neco_a_01199.
[28] R. Dey, F.M. Salem, Gate-variants of gated recurrent unit (GRU) neural networks, in: 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), IEEE, 2017, pp. 1597–1600.
[29] M. Alhussein, K. Aurangzeb, S.I. Haider, Hybrid CNN-LSTM model for short-term individual household load forecasting, IEEE Access 8 (2020) 180544–180557, https://doi.org/10.1109/ACCESS.2020.3028281.
[30] D.A. Iancu, N. Trichakis, Pareto efficiency in robust optimization, Manag. Sci. 60 (1) (2014) 130–147, https://doi.org/10.1287/mnsc.2013.1753.
[31] T. Jia, J. Zhou, X. Liu, A Daily Power Generation Optimized Operation Method of Hydropower Stations with the Navigation Demands Considered, MATEC Web of Conferences, vol. 246, EDP Sciences, 2018, p. 01065.