

# A varying comprehensive hydropower coefficient for medium/long-term operation of a single reservoir

Wenting Gong, Pan Liu, Lei Cheng, He Li and Zhikai Yang

## ABSTRACT

In medium/long-term reservoir operation, the hydropower output is calculated from  $k \times q \times h$ , where  $q$  is the power discharge,  $h$  is the water head, and  $k$  is the comprehensive hydropower coefficient.  $k$  indicates the conversion efficiency from water power to electricity, however, it is standard practice to use a constant  $k$ . We developed a novel method to derive the varying  $k$  based on observed big data. The operational frequencies of different units for time (multiple periods) and space (multiple units) were accounted based on the observed big data from each unit, and then weights were obtained. Finally,  $k$  was derived by integrating the efficiency curves (hill charts) of the different units with their weights. The Three Gorges Project, China, was selected for a case study. Results indicated that: (1) the varying  $k$  value can improve hydropower simulation accuracy, (2) simulations using 10-day intervals have a higher accuracy for hydropower calculation than daily and monthly scales, (3) the evaluation of hydropower plant benefits is sensitive to  $k$ , and there is potential for producing more hydropower. These findings are highly relevant to the operation of hydropower plants and to the evaluation of medium/long-term hydropower generation for a hydropower plant.

**Key words** | comprehensive hydropower coefficient, hydropower benefit evaluation, unit efficiency curve

**Wenting Gong**  
**Pan Liu** (corresponding author)  
**Lei Cheng**  
**He Li**  
**Zhikai Yang**  
 State Key Laboratory of Water Resources and  
 Hydropower Engineering Science,  
 Wuhan University,  
 Wuhan 430072,  
 China  
 and  
 Hubei Provincial Collaborative Innovation Center  
 for Water Resources Security,  
 Wuhan 430072,  
 China  
 E-mail: liupan@whu.edu.cn

## HIGHLIGHTS

- A temporal-spatial aggregation method that uses big data for determining the varying comprehensive hydropower coefficient.
- The derived varying comprehensive coefficient improves the hydropower energy production simulation accuracy.
- The derived varying comprehensive coefficient is sensitive to evaluate the hydropower plant benefits.

## NOTATION

The following symbols are used in this paper:

$A$  = total output numbers for the entire plant;  
 $A_i$  = output percent in interval  $i$  (%);

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY-NC-ND 4.0), which permits copying and redistribution for non-commercial purposes with no derivatives, provided the original work is properly cited (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

doi: 10.2166/nh.2020.158

$a_i$  = output numbers in interval  $i$ ;  
 $B$  = total output numbers of the plant in the interval  $i$ ;  
 $B_j$  = output percent in the  $j$  category (%);  
 $b_j$  = total output numbers of the  $j$  category;  
 $E_o$  = actual power generation (billion kWh);  
 $E_k$  = benchmark power generation (billion kWh);  
 $H$  = water head (m);  
 $i$  = index of the output intervals;

$I_m$  = inflow in the  $m$  period ( $\text{m}^3/\text{s}$ );  
 $j$  = index of the unit categories;  
 $k$  = comprehensive hydropower coefficient for the plant;  
 $k_{ij}$  = varying values of the  $j$  type unit in the  $i$  interval;  
 $M$  = output levels;  
 $m$  = numbers of periods;  
 $N$  = output (kW);  
 $P$  = unit categories;  
 $P_m$  = output in the  $m$  period (kW);  
 $PU_m$  = upper limits of the output (kW);  
 $PL_m$  = lower limits of the output (kW);  
 $Q_m$  = outflow in the  $m$  period ( $\text{m}^3/\text{s}$ );  
 $QU_m$  = upper limits of the outflow ( $\text{m}^3/\text{s}$ );  
 $QL_m$  = lower limits of the outflow ( $\text{m}^3/\text{s}$ );  
 $q$  = power discharge ( $\text{m}^3/\text{s}$ );  
 $V_m$  = storage capacity in the  $m$  period ( $\text{m}^3$ );  
 $VU_m$  = upper limits of the storage capacity ( $\text{m}^3$ );  
 $VL_m$  = lower limits of the storage capacity ( $\text{m}^3$ );

## INTRODUCTION

Due to the limitations of traditional fossil energy (Ming et al. 2017), renewable energy is playing an increasingly important role in energy supply. Water power is a typical example of renewable energy, especially in China. There are many studies on improving the efficiency of hydroelectric power production to increase the availability of energy (Inglesi-Lotz & Blignaut 2014), and reservoir operation is a key issue for improving hydropower efficiency. A large number of these studies have focused on the optimization method for reservoir operation (Liao et al. 2017; Feng et al. 2018), while the hydropower efficiency evaluation is seldom studied (Chang et al. 2017).

In medium/long-term reservoir operation (Uen et al. 2018), hydropower output  $N$  is calculated by  $k \times q \times h$ , where  $q$  is the power discharge,  $h$  is the water head, and  $k$  is the comprehensive hydropower coefficient, which indicates the conversion efficiency from water power to electricity (Hidalgo et al. 2009). The coefficient  $k$  equals the gravitational constant multiplied by the efficiency of the turbine-generator set (Hidalgo et al. 2014), and is a key parameter in the simulation of hydropower generation. The  $k$  value is usually taken as a fixed constant for the

entire plant to simplify calculations (Hidalgo et al. 2012). However, the  $k$  value varies among different units, and with the water head and hydropower discharge (Liu et al. 2012).

Much research has been done on identifying the factors that determine  $k$ . For example, Liu et al. (2012) identified influencing factors and Xu et al. (2017) did a sensitivity analysis on units that affect the  $k$  value. Diniz et al. (2007) and Cordova et al. (2014) studied the hydraulic efficiency of turbines and established a relationship between efficiency and  $h$  and  $q$  (Finardi & da Silva 2006; Diniz et al. 2007; Cordova et al. 2014). In these methods, the relationship between  $k$  and  $h$ , and  $k$  and  $q$  were based on direct estimation from observed data. However, an average value of  $h$  does not meaningfully describe the variation of hydropower plants. Therefore, these methods are restricted in applicability to short-term operations, and are difficult to be used for medium/long-term operations.

Other studies have focused on adjusting the efficiency curves or exploring the performances of the units. For example, Barros & Peypoch (2007) adopted a random cost frontier method to analyze technical efficiency. Barros (2008) sought to identify the best practices for improving the performance of hydroelectric generating plants. Hidalgo et al. (2010) used a simulator to improve the efficiency and reliability of plant data. Cordova et al. (2014) described a system to calculate the efficiency of each individual unit, while Li et al. (2014) described a three-dimensional interpolation technique to represent the generation function of each individual unit. Hidalgo et al. (2014) used an iterative calculation to obtain the efficiency functions. However, the improvement in accuracy of the efficiency curves was used for a single unit type, and could not be applied to the entire hydropower plant operation.

Advances in technology and communication have led to big data research. Big data form a good basis for overcoming the limitations of existing methods, and for investigating the variation rules for  $k$ . Therefore, the purpose of this study is to develop a novel method for obtaining the variations in  $k$  using observed big data. Our intention is that the proposed method should reduce the generation simulation and efficiency evaluation errors for medium/long-term and short-term operations. Our specific aims are to: (1) determine a comprehensive varying  $k$  using a temporal-spatial

aggregation estimation method, (2) test the derived  $k$  using the observed big data, and (3) re-evaluate reservoir operation benefits using the derived  $k$ . The rest of the paper is organized as follows. The ‘Methodology’ presents two estimation methods to generate the  $k$ – $h$  relationship for an entire hydropower plant. ‘Case study’ describes a case study using the Three Gorges Project (TGP), China. In ‘Results and discussion’, the varying  $k$  value is tested for different time periods, and its impacts on reservoir operation efficiency are evaluated. ‘Conclusions’ summarizes the conclusions of this study.

## METHODOLOGY

Two estimation methods for determining a varying  $k$  (the direct method and the temporal-spatial aggregation method) are presented for comparative purposes (Figure 1). The direct method is carried out using observed output, water head, and discharge data (section: ‘The direct estimated method’), while the temporal-spatial aggregation

method built the  $k$ – $h$  relationship using the efficiency curve (it is also called hill chart) and the observed big data (section: ‘The temporal-spatial aggregation estimated method’). These two methods are then tested for hydropower simulation using an error analysis.

### Derivation of varying $k$

#### The direct estimated method

Existing studies commonly use observed  $N$ ,  $q$ , and  $h$  values to calculate the  $k$  value. The value of  $k$  can be estimated (Hidalgo *et al.* 2014) using the following equation:

$$k = \frac{N}{q \times h} \quad (1)$$

However, the water head varies with time, and in the medium to long term an average value is unable to meaningfully describe the processes of the hydropower plant. Thus, the  $k$  value derived using Equation (1) is not always suitable

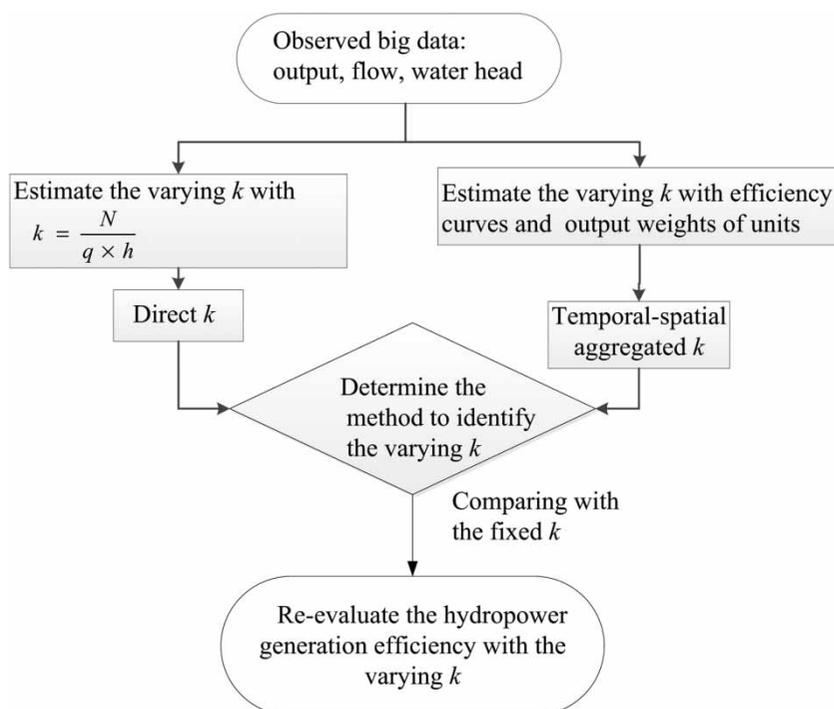


Figure 1 | Flow chart for the estimation of varying  $k$ .

since the value of  $h$  is difficult to be determined. Therefore, different  $k$  values should be derived at different time scales, respectively.

### The temporal-spatial aggregation estimated method

The efficiency curves of plant units are used to improve the estimation accuracy of  $k$ . The  $k$  value is obtained by interpolation from efficiency curves when  $h$  and  $q$  are known.

The hydropower output of each unit is divided into several ( $M$ ) levels from minimum to maximum output, thus forming  $M - 1$  intervals. The output for the entire plant in each interval is counted and the frequency  $A_i$  was determined as follows:

$$A_i = \frac{a_i}{A} \quad (2)$$

where  $a_i$  is the output numbers in interval  $i$ , and  $A$  represents the total output numbers for the entire plant. The output percent  $A_i$  ( $i = 1, 2, \dots, M - 1$ ) reflects the operation time for the  $i$  interval, and can be used as a weight to combine the operation time for each unit.

Similarly, the output percent in each interval for variable unit types is obtained to reflect the combination of different units ( $P$ ). The output numbers of each unit are accounted and the frequency  $B_j$  was determined as follows:

$$B_j = \frac{b_j}{B} \quad (3)$$

where  $b_j$  is the total output numbers of the  $j$  category, and  $B$  is the total output numbers of the plant in the interval  $i$ .  $B_j$  ( $j = 1, 2, \dots, P$ ) is the output percent of the  $j$  category, and is used as weights in the spatial aggregation among different units.

After multiplying the  $k_{ij}$  value with the above weights, the final  $k$  can be calculated by:

$$k = \sum k_{ij} \times A_i \times B_j \quad (4)$$

where  $k_{ij}$  is the varying value of the  $j$  type unit in the  $i$  interval, and was obtained using its efficiency curve. Since the derivation used minute-level data, the  $k$  value is suitable for any time scales.

### Evaluation indexes for the $k$ value

An estimation is implemented to test the derived  $k$ - $h$  relationship. The estimation is obtained for the direct and for the temporal-spatial aggregation methods. The accuracy of the varying coefficient is evaluated by comparing the estimated results with the observed power output. The power generation, the error percentage of the power generation, the mean error percentage of 10 days, and the root mean square error (*RMSE*) are chosen as the evaluation criteria.

### Efficiency evaluation of hydropower generation

The relative hydropower utilization rate (*RHUR*) is used to evaluate the efficiency of the actual operation as follows (Chang et al. 2017):

$$RHUR = \frac{E_0 - E_k}{E_k} \times 100\% \quad (5)$$

where  $E_0$  represents actual power generation, and  $E_k$  represents benchmark power generation. If  $RHUR > 0$ , then the operation for a plant is acceptable because the produced hydropower energy is greater than the benchmarking criterion.

## CASE STUDY

### The TGP

The TGP is, currently, the largest multipurpose hydro-development project ever built. The TGP is crucial for the water resources development of China's largest river, the Yangtze River (Figure 2). The benefits provided by the TGP include flood control, power generation, and navigation improvement. The TGP receives inflow from a  $4.5 \times 10^3$  km long channel with a contributing drainage area of  $1 \times 10^6$  km<sup>2</sup>. The mean annual runoff at the dam site is  $4.51 \times 10^{11}$  m<sup>3</sup>. The flood storage capacity of the TGP is very large ( $2.215 \times 10^{10}$  m<sup>3</sup>), and plays a very important role in the flood control of the Yangtze River. The TGP has 32 sets of  $70 \times 10^4$  kW hydraulic turbo generators, which equates to  $22.50 \times 10^4$  kW in total installed capacity. The TGP

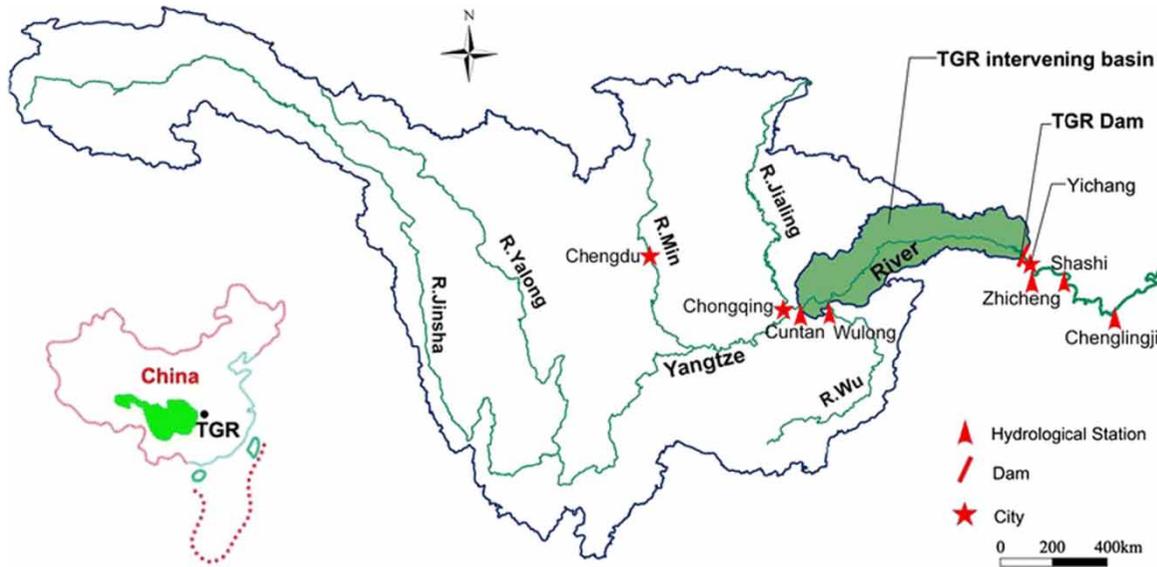


Figure 2 | TGP Hydropower Plant.

produces an annual electricity output of  $8.468 \times 10^{10}$  kWh, and a large proportion of this is used to supply eastern and central China. The TGP also improves the navigation conditions downstream in the dry season, as well as in the 660 km-long waterway upstream, and improves downstream water quality during the dry season. In addition, the TGP enhances fish habitats in the reservoir, as well as enhancing tourism and recreational activities (Liu *et al.* 2011). Table 1 lists the TGP parameters (Li *et al.* 2010).

Although the hydropower units have the same rated power, they are produced by four different manufacturers. Based on the manufacturer and the time of installation (Table 2), the units were classified into eight categories, each with a slight difference of performance.

Table 1 | Parameters of the Three Gorges reservoir

| Parameter                       | Value | Unit                   |
|---------------------------------|-------|------------------------|
| Normal pool level               | 175   | m                      |
| Flood limited water level       | 145   | m                      |
| Dry control water level         | 155   | m                      |
| Flood protection storage        | 22.15 | billion m <sup>3</sup> |
| Conservation regulating storage | 16.5  | billion m <sup>3</sup> |
| The total reservoir storage     | 39.3  | billion m <sup>3</sup> |

Table 2 | Unit categories for the Three Gorges power station

| Factory      | Location                              | Unit number              |
|--------------|---------------------------------------|--------------------------|
| VGS          | Left bank                             | 1–3, 7–9                 |
| ALSTOM       | Left bank, right bank and underground | 4–6, 10–14, 19–22, 29–30 |
| Orient Motor | Right bank and underground            | 15–18, 27–28             |
| Harbin Motor | Right bank and underground            | 23–26, 31–32             |

## Data description

### Observed big data

Observed data from 2015 to 2018 were used to illustrate the method. The observed data include: water levels in the upstream and downstream of the reservoir at 5 minute intervals, hydropower for each unit at 5 minute intervals, and the power flow of the entire plant every 2 hours. 1,280,000 observed data were used to derive the varying  $k$  value. This is a typical example of combining big data with a case study, and thus obtaining more reliable results.

The data during 2015–2016 were used to estimate the varying  $k$ , while the data during 2017–2018 were used to test the estimation accuracy by comparing the estimated and observed hydropower generation.

## Operation parameters

The efficiency curves of the units were used to estimate  $k$  (Figure 3).

In addition, the maximum output of the entire plant, the relationship between water levels and reservoir capacity, and the relationship between release and downstream water levels were used in the reservoir simulation and optimization.

## RESULTS AND DISCUSSION

### Derivation results of varying $k$

#### The direct estimated results

The varying  $k$  was obtained by applying the direct method to the 2015–2016 data. Due to the characteristic of medium/long term operation, 10 days is selected as the time scale for the calculation, the average value of  $N$ ,  $q$ , and  $h$  within 10 days is solved, and the corresponding  $k$  value is obtained according to the formula; the 72 points are shown in Figure 4. It is clear from the calculated results that there is a significant relationship between  $k$  and  $h$ . As a result, a function curve was used to fit the relationship between  $k$  and  $h$  as shown in Figure 4. According to the water head of the power station, the corresponding  $k$

value can be obtained by the fitting curve, and then the power generations are estimated.

However, the correlation is not perfect, and this fitting may decrease the estimation accuracy. Furthermore, the average water heads were used instead of the actual water heads, and this led to errors in the  $k$  value.

### The temporal-spatial aggregation estimated results

Each unit's output ranges from 0 to  $70 \times 10^4$  kW. However, in actual operations the output was seldom less than  $50 \times 10^4$  kW. Therefore, output ranges from  $50 \times 10^4$  kW to  $70 \times 10^4$  kW were used.

There were a total of five levels, and these formed four intervals. Using the observed data from 2015 to 2016, the actual output operation times for the different intervals were counted and the frequencies were used as weights (Table 3). Then the weights of the eight categories are counted according to the observed data, as shown in Table 4.

For the  $i$  interval, the value of  $k_{ij}$  was calculated using the efficiency curve, as shown in Figure 5. The  $k_i$  value was calculated using  $k_{ij}$  and  $b_j$  for interval  $i$ , as shown in Figure 6. The varying  $k$  value was then obtained using  $k_i$  and  $a_i$  for the entire plant.

Finally, the relationship between  $k$  and  $h$  was derived (Figure 7). According to the observation data, the water head at the plant was 76 m–110 m. Thus, the  $k$ – $h$  relationship was estimated only for this range. It is observed that a

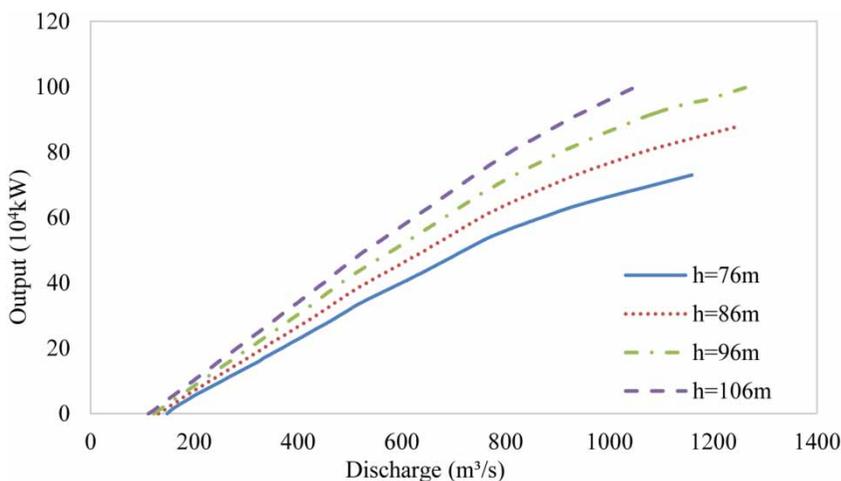


Figure 3 | Schematic efficiency curve for the ALSTOM units of Three Gorges Hydropower Plant.

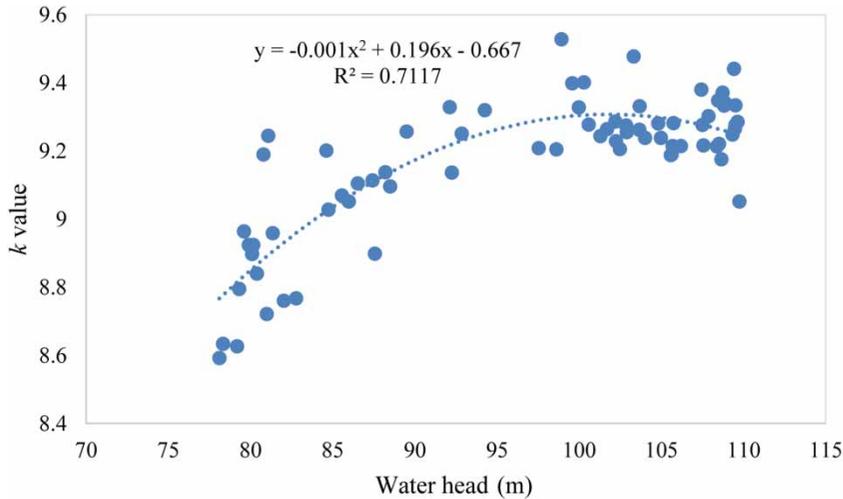


Figure 4 | Direct relationship between  $k$  and  $h$ .

Table 3 | Weights  $a_i$  of  $k_{ij}$  for the Three Gorges hydropower plant

| Output range ( $10^4$ kW) | [50–55) | [55–60) | [60–65) | [65–70) |
|---------------------------|---------|---------|---------|---------|
| Weight                    | 0.014   | 0.047   | 0.256   | 0.683   |

water head of 94 m is a critical point. The value of  $k$  increases with increasing water head before this point, and then decreases after this point.

### Test results of varying $k$

#### Estimation using the temporal-spatial aggregation method

The  $k$  value can be determined based on the observed water head. Combining with observed power discharge, the

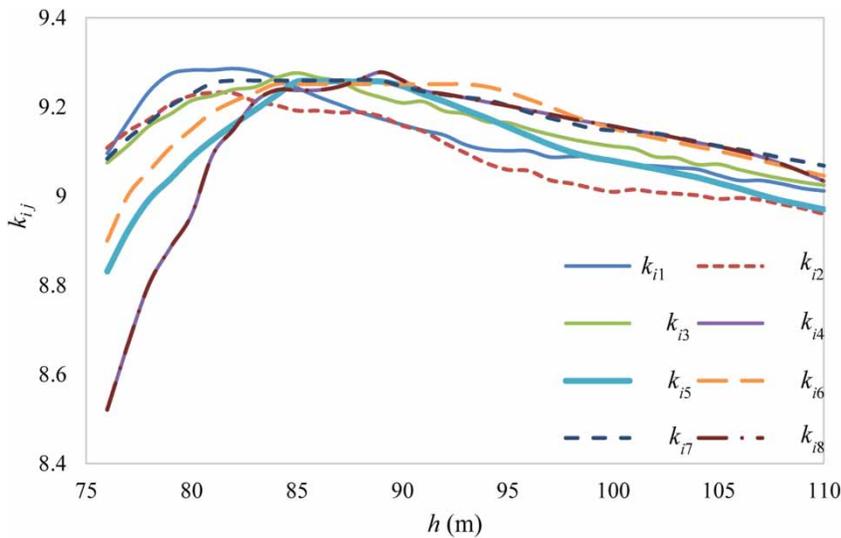
estimated outputs can be obtained. Ten days (a traditional Chinese time period) was chosen as the time interval in the computation.

The estimated results for 2017–2018 were compared with the observed power generation. Differences between estimated and observed values are shown in Table 5. An estimation using a constant  $k$  value has been included to demonstrate the inherent characteristic of  $k$ , and to highlight the necessity of using a varying  $k$  value.

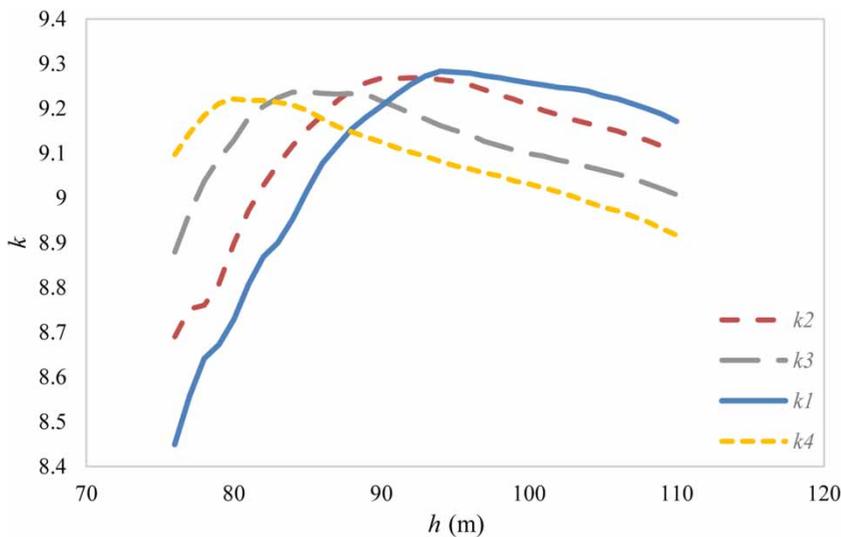
It is clear that the accuracy of the temporal-spatial aggregated  $k$  is higher than that of the fixed  $k$  and the direct  $k$ . As shown in Table 5, the annual power generation errors for the temporal-spatial aggregated  $k$  are only slightly greater than 0.7%, but the generation errors for the fixed  $k$  exceed 1.9%. In addition, the *RMSE* of the temporal-spatial

Table 4 | Weights  $b_j$  of  $k_{ij}$  for the Three Gorges hydropower plant

| Unit Number | Output range ( $10^4$ kW) | Weight |
|-------------|---------------------------|--------|---------------------------|--------|---------------------------|--------|---------------------------|--------|
| 1–3, 7–9    | [50–55)                   | 0.032  | [55–60)                   | 0.100  | [60–65)                   | 0.096  | [65–70)                   | 0.272  |
| 4–6, 10–14  |                           | 0.088  |                           | 0.223  |                           | 0.129  |                           | 0.304  |
| 19–22       |                           | 0.196  |                           | 0.139  |                           | 0.232  |                           | 0.106  |
| 15–18       |                           | 0.390  |                           | 0.221  |                           | 0.160  |                           | 0.092  |
| 23–26       |                           | 0.154  |                           | 0.128  |                           | 0.216  |                           | 0.108  |
| 29–30       |                           | 0.035  |                           | 0.076  |                           | 0.057  |                           | 0.038  |
| 27–28       |                           | 0.014  |                           | 0.043  |                           | 0.061  |                           | 0.043  |
| 31–32       |                           | 0.091  |                           | 0.070  |                           | 0.049  |                           | 0.037  |



**Figure 5** | Values of  $k_{ij}$  versus water head.



**Figure 6** | Values of  $k_j$  versus water head.

aggregated  $k$  is smaller than that of the fixed  $k$ , indicating a good estimation.

The accuracy of the direct estimation method is the lowest of the three methods, and the error percentages for the entire year and for the mean of each 10 days were both almost 2.5%. The *RMSE* for the direct estimation method is much higher than for the other estimation methods. Thus, it can be concluded that the directly determined  $k$  is not the most suitable for power generation

simulation, and therefore temporal-spatial aggregation was chosen as the estimation method for the varying  $k$ .

As shown in Figures 8 and 9, the relative error has huge differences in the flood period, thus the estimated accuracy of the  $k$  value is lower than in the non-flood period. On the contrary, the direct estimation of the  $k$  value in the flood season is much better than the other two estimation methods. It can be considered to combine the temporal-spatial aggregated method and the direct-determined

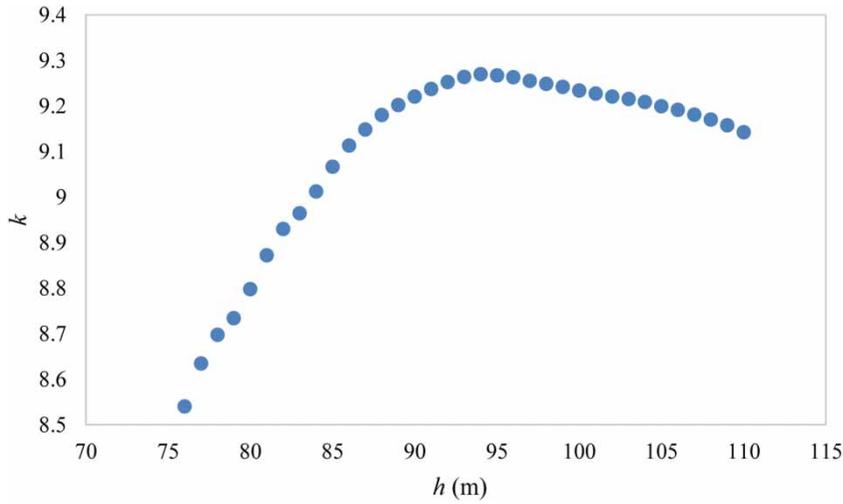


Figure 7 | The obtained  $k$ - $h$  relationship by using temporal-spatial aggregation method.

Table 5 | Estimated results of power generation in 2017–2018

| Year | Observed power generation (billion kWh) | Estimated condition             | Estimated power generation (billion kWh) | Power generation error percentage (%) | Mean of every 10 days' error percentage (%) | RMSE |
|------|---|---------------------------------|--|---------------------------------------|---|------|
| 2017 | 97.61                                   | $k = 9.0$                       | 95.76                                    | - 1.89                                | 1.45  | 1.64 |
|      |   | Direct $k$                      | 93.76                                    | - 3.94                                | 2.45  | 2.53 |
|      |   | Temporal-spatial aggregated $k$ | 96.89                                    | - 0.74                                | 0.96  | 1.03 |
| 2018 | 101.62                                  | $k = 9.0$                       | 100.33                                   | - 1.27                                | 1.61  | 1.85 |
|      |   | Direct $k$                      | 98.10                                    | - 3.46                                | 2.18  | 2.34 |
|      |   | Temporal-spatial aggregated $k$ | 101.44                                   | - 0.18                                | 1.30  | 1.53 |

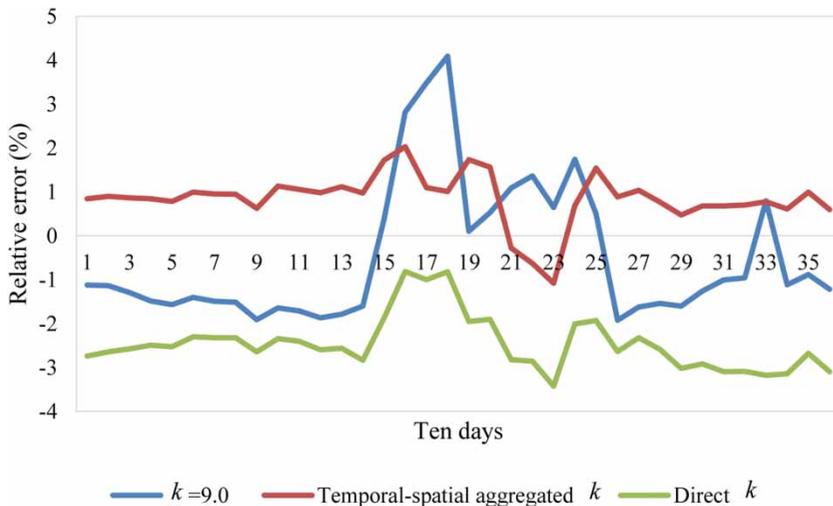
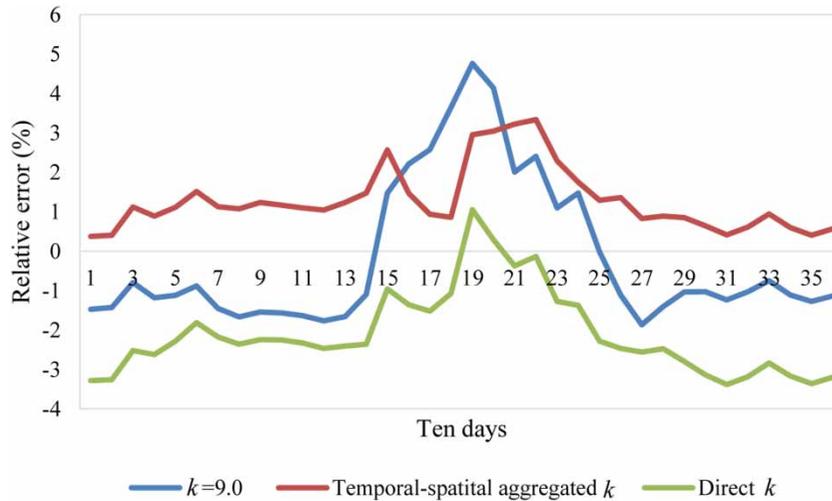


Figure 8 | The errors in power generation for 10-day periods in 2017.



**Figure 9** | The errors in power generation for 10-day periods in 2018.

method during the year to further improve the estimation accuracy of the  $k$  value.

### Estimation for different time intervals

The  $k$  varies constantly according to what has been described above. Therefore, it is necessary to investigate the relationship between the benefits' evaluation and the time interval, and to determine whether a shorter interval results in a better estimation. It is also necessary to investigate the stability of the proposed method under different conditions. Repeated estimations were carried out for three time intervals: 1 day, 10 days, and 1 month.

Comparing with the observed data, the estimated annual power generation was almost the same for all three time intervals (Table 6). The mean error percentage for daily intervals was the largest. These results ran contrary to

expectations. The estimations for monthly intervals yielded a reasonable mean error percentage and a variance that was the highest of the three intervals.

It is concluded that 10-day intervals are optimal for the benefits' evaluation process. Both the estimated errors and the details are reasonable. Monthly intervals are useful when the operating period is long. Daily intervals are efficient when a high evaluation accuracy is required. Different time intervals could be chosen based on the concrete demand.

### Conventional operation benefit evaluation

As shown in Figure 10, the conventional operation of the Three Gorges Power Station is implemented using reservoir operating rule curves (Feng et al. 2017).

Table 7 lists a comparison of two  $k$  value approaches for hydropower generation based on the observed data.

**Table 6** | Estimated results for different time intervals in 2017–2018

| Year | Time interval | Estimated power generation (billion kWh) | Power generation error percentage (%) | Mean of error percentage (%) | Mean of variance for error percentage |
|------|---------------|--|---------------------------------------|------------------------------|---------------------------------------|
| 2017 | Day           | 98.46                                    | 0.87                                  | 0.98                         | 1.31                                  |
|      | 10 days       | 96.89                                    | -0.74                                 | 0.96                         | 1.06                                  |
|      | Month         | 97.09                                    | -0.53                                 | 1.07                         | 1.38                                  |
| 2018 | Day           | 103.15                                   | 1.51                                  | 1.24                         | 2.21                                  |
|      | 10 days       | 101.44                                   | -0.17                                 | 1.30                         | 2.33                                  |
|      | Month         | 101.64                                   | 0.02                                  | 1.46                         | 2.97                                  |

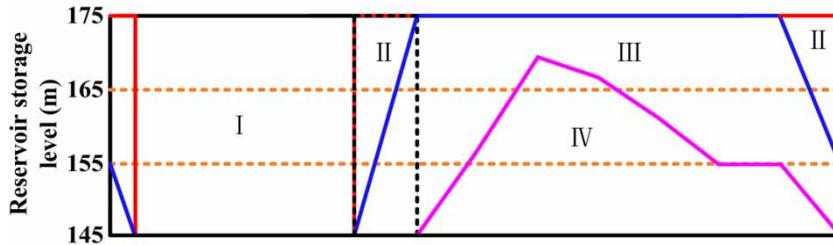


Figure 10 | Conventional operation rule curves for the Three Gorges Project.

Table 7 | The results of conventional operations in 2017 and 2018

| Year | <i>k</i> value                       | Simulated power generation (billion kWh) | RHUR (%) |
|------|--------------------------------------|--|----------|
| 2017 | <i>k</i> = 9.0                       | 92.31                                    | 5.74     |
|      | Temporal-spatial aggregated <i>k</i> | 91.81                                    | 6.33     |
| 2018 | <i>k</i> = 9.0                       | 93.26                                    | 8.96     |
|      | Temporal-spatial aggregated <i>k</i> | 93.49                                    | 8.70     |

When the *k* value is a fixed constant of 9.0, the calculated energy produced in 2017 and 2018 is 92.31 and 93.26 billion kWh, respectively. When a varying *k* value is used, the calculated energy produced is 91.81 and 93.49 billion kWh, respectively. The observed power generation in 2017 and 2018 was 97.61 billion kWh and 101.62 billion kWh, respectively. It is indicated that the actual operation is superior to the conventional rules by 6%–8%. However, the RHUR for the varying *k* is different to the fixed *k*. It is indicated that the improvement in benefit should be different to that obtained in the fixed evaluation. The varying *k* produces a more realistic hydropower generation by improving the estimation accuracy.

## Potential benefit evaluation

### Optimal operation model

The optimal reservoir operation can be used as a benchmark to evaluate the potential benefit of hydropower generation. The optimal reservoir operation model is built with the single objective function of maximizing hydropower generation. The dynamic programming algorithm was used to

solve the optimization model based on the historical inflows and the results were compared with the observed data. Moreover, the following constraints were considered in optimal operation:

1. Water balance constraint:

$$V_{m+1} = V_m + (I_m - Q_m) \cdot t \quad (6)$$

where *m* is the number of periods,  $V_m$  is the storage capacity in the *m* period, and  $I_m$  and  $Q_m$  are the inflow and outflow of the hydropower station, respectively.

2. Storage capacity constraints:

$$VL_m \leq V_m \leq VU_m \quad (7)$$

where  $VL_m$  and  $VU_m$  are the lower and upper limits of the storage capacity of the hydropower station, respectively.

3. Release constraints:

$$QL_m \leq Q_m \leq QU_m \quad (8)$$

where  $QL_m$  and  $QU_m$  are the lower and upper limits of the outflow of the hydropower station, respectively.

4. Output constraints:

$$PL_m \leq P_m \leq PU_m \quad (9)$$

where  $PL_m$  and  $PU_m$  are the lower and upper limits of the output of the hydropower station, respectively.

5. Non-negative constraints and other constraints.

### Optimal operation within different boundaries of water levels

Since the actual water level is higher than the flood-limited water level (the conventional upper water level boundary) during flood seasons (Liu et al. 2015), two water level boundaries were available as constraints for evaluating potential hydropower benefits. According to the design requirements of the power station, in the non-flood season, the upper and lower water levels are 175 m and 145 m, respectively. In the flood season, the feasible interval ranges from 140 m to 145 m. The optimal operation results for 2017 and 2018 were obtained using deterministic dynamic programming.

The results of the optimal model with different upper water level limits are shown in Table 8. It can be found that the corresponding  $RHUR$  values are negative based on the observed data, indicating that the actual operation does not produce maximum power, and that there is considerable potential for increasing the hydropower generated. The results of the varying  $k$  are about 1.7% higher than the results of the fixed one, which indicates that the varying  $k$  can accurately evaluate the optimal space of operation benefits and the optimal operation is sensitive to the  $k$  value.

Table 8 also lists the optimal results within actual operation boundary of water levels. Due to the increased water levels, the power generation potential both noticeably increased in new intervals for 2017 and 2018 than the designed one. In 2017 and 2018, the benefits' rate with the actual operation water level was 0.72% and 0.65% higher than for the designed operation water level for the varying simulation. This clearly illustrates that potential hydropower generation is higher with different operation water level boundaries.

### CONCLUSIONS

This study proposed the temporal-spatial aggregation method for estimating varying  $k$  to minimize the errors in hydropower generation simulation for medium/long-term operations. The  $k$ - $h$  relationship was obtained using unit efficiency curves and observed big data. The TGP was used as a study case. It was concluded that:

1. The method proposed was successfully used to determine the varying  $k$  for the entire power plant for medium/long-term reservoir operations. The temporal-spatial aggregation method produced better results than the direct estimation of  $k$  method, and is suitable for calculations in any time scales.
2. Using the varying  $k$  value improved the simulation accuracy. Using a time interval of 10 days resulted in the relative highest accuracy in hydropower generation simulation.
3. The derived  $k$  plays an important role in benefit evaluation of reservoir operation. Conventional power generation operations can be improved by using the varying  $k$ . The potential for improvement is large.

The method proposed can be applied to other hydropower plants to improve hydropower simulation and to provide a more balanced benefit evaluation. However, further research is needed to assess measurement errors.

### Data availability section

All data used during the study are proprietary and may only be provided with restrictions (e.g., anonymized data). Including the observed big data, the operation parameters, the maximum output of the entire plant, the relationship

**Table 8** | Optimal operation in 2017 and 2018

| Year | $K$ value                       | Simulated power generation with designed boundaries (billion kWh) | $RHUR$ (%) | Simulated power generation with historical boundaries (billion kWh) | $RHUR$ (%) |
|------|---------------------------------|---|------------|---|------------|
| 2017 | $k = 9.0$                       | 102.76  | - 5.01     | 103.41  | - 5.61     |
|      | Temporal-spatial aggregated $k$ | 104.84  | - 6.70     | 105.43  | - 7.42     |
| 2018 | $k = 9.0$                       | 102.04  | - 0.41     | 102.71  | - 1.06     |
|      | Temporal-spatial aggregated $k$ | 103.85  | - 2.15     | 104.55  | - 2.80     |

between water levels and reservoir capacity, and the relationship between outflow and downstream water levels were used in the reservoir simulation and optimization. All of the data are illustrated in the section ‘Data description’.

## ACKNOWLEDGEMENTS

This study was supported by the National Key Research and Development Program of China (2016YFC0402202), the National Natural Science Foundation of China (U1865201) and the Innovative Research Groups of the Natural Science Foundation of Hubei, China (2017CFA015). The authors thank the editor and anonymous reviewers for their comments that improved the paper. The authors declare that they have no conflict of interest.

## REFERENCES

- Barros, C. P. 2008 Efficiency analysis of hydroelectric generating plants: a case study for Portugal. *Energy Economics* **30**, 17.
- Barros, C. P. & Peypoch, N. 2007 The determinants of cost efficiency of hydroelectric generating plants: a random frontier approach. *Energy Policy* **35**, 4463–4470.
- Chang, J., Li, Y., Yuan, M. & Wang, Y. 2017 Efficiency evaluation of hydropower station operation: a case study of Longyangxia station in the Yellow River, China. *Energy Economics* **135**, 9.
- Cordova, M. M., Finardi, E. C., Ribas, F. A. C., de Matos, V. L. & Scuzziato, M. R. 2014 Performance evaluation and energy production optimization in the real-time operation of hydropower plants. *Electric Power Systems Research* **116**, 201–207.
- Diniz, A. L., Esteves, P. P. I. & Sagastizábal, C. A. 2007 A mathematical model for the efficiency curves of hydroelectric units. In: *2007 IEEE Power Engineering Society General Meeting*, Tampa, FL, USA.
- Feng, M., Liu, P., Guo, S., Shi, L., Deng, C. & Ming, B. 2017 Deriving adaptive operating rules of hydropower reservoirs using time-varying parameters generated by the EnKF. *Water Resources Research* **53**, 6885–6907.
- Feng, Z.-k., Niu, W.-j., Cheng, C.-t. & Wu, X.-y. 2018 Peak operation of hydropower system with parallel technique and progressive optimality algorithm. *International Journal of Electrical Power & Energy Systems* **94**, 267–275.
- Finardi, E. C. & da Silva, E. L. 2006 Solving the hydro unit commitment problem via dual decomposition. *IEEE Transactions on Power Systems* **21**, 10.
- Hidalgo, I. G., Soares, S., Fontane, D. G. & Cicogna, M. A. 2009 Management and analysis of data from hydroelectric plants. In *Power Systems Conference and Exposition*, Seattle, WA, USA.
- Hidalgo, I. G., Fontane, D. G., Soares, S., Cicogna, M. A. & Lopes, J. E. G. 2010 Data consolidation from hydroelectric plants. *Journal of Energy Engineering* **136**, 8.
- Hidalgo, I. G., Fontane, D. G., Arabi, M., Lopes, J. E. G., Andrade, J. G. P. & Ribeiro, L. C. A. J. 2012 Evaluation of optimization algorithms to adjust efficiency curves for hydroelectric generating units. *Energy Engineering* **138** (4), 7.
- Hidalgo, I. G., Fontane, D. G., Lopes, J. E. G., Andrade, J. G. P. & de Angelis, A. F. 2014 Efficiency curves for hydroelectric generating units. *Water Resources Planning and Management* **140** (1), 6.
- Inglesi-Lotz, R. & Blignaut, J. N. 2014 Improving the electricity efficiency in South Africa through a benchmark-and-trade system. *Renewable and Sustainable Energy Reviews* **30**, 833–840.
- Li, X., Guo, S., Liu, P. & Chen, G. 2010 Dynamic control of flood limited water level for reservoir operation by considering inflow uncertainty. *Journal of Hydrology* **391**, 124–132.
- Li, X., Li, T., Wei, J., Wang, G. & Yeh, W.-G. 2014 Hydro unit commitment via mixed integer linear programming. *IEEE Transactions on Power Systems* **29**, 10.
- Liao, S.-l., Liu, B.-x., Cheng, C.-t., Li, Z.-f. & Wu, X.-y. 2017 Long-term generation scheduling of hydropower system using multi-core parallelization of particle swarm optimization. *Water Resources Management* **31**, 2791–2807.
- Liu, P., Cai, X. & Guo, S. 2011 Deriving multiple near-optimal solutions to deterministic reservoir operation problems. *Water Resources Research* **47**, W08506. doi:10.1029/2011WR010998.
- Liu, R., Wei, J. & Xiang, L. 2012 Calculation of comprehensive power output coefficient of a hydropower station and its effects on simulating the regulation process. *South-to-North Water Diversion and Water Science & Technology* **10**, 4.
- Liu, P., Li, L., Guo, S., Xiong, L., Zhang, W., Zhang, J. & Xu, C. 2015 Optimal design of seasonal flood limited water levels and its application for the Three Gorges Reservoir. *Journal of Hydrology* **527**, 1045–1053.
- Ming, B., Liu, P., Guo, S., Zhang, X., Feng, M. & Wang, X. 2017 Optimizing utility-scale photovoltaic power generation for integration into a hydropower reservoir by incorporating long- and short-term operational decisions. *Applied Energy* **204**, 432–445.
- Uen, T. S., Chang, F. J., Zhou, Y. & Tsai, W. P. 2018 Exploring synergistic benefits of Water-Food-Energy Nexus through multi-objective reservoir optimization schemes. *Science of the Total Environment* **633**, 341–351.
- Xu, B., Chen, D., Gao, X. & Wu, C. 2017 Sensitivity analysis of a Pelton hydropower station based on a novel approach of turbine torque. *Energy Conversion and Management* **148**, 785–800.