Impact of the Data Quality from Hydroelectric Plants in the Past Operation Analysis using a Middle Term Simulation Tool

Ieda G. Hidalgo¹, Secundino Soares F.², Darrell G. Fontane³, Marcelo A. Cicogna⁴ and João E. G. Lopes⁵

¹ Ph.D. Electrical Engineering, Department of Electrical and Computer Engineering, State University of Campinas (UNICAMP) – Brazil; e-mail: iedahidalgo@gmail.com
² Professor, Department of Electrical and Computer Engineering, State University of Campinas (UNICAMP) – Brazil; e-mail: dino@cose.fee.unicamp.br
³ Professor, Department of Civil and Environment Engineering, Colorado State University (CSU) – EUA; e-mail: darrell.fontane@colostate.edu
⁴ Ph.D. Electrical Engineering, Anhanguera Educacional (AESA) – Brazil; e-mail: marcelocicogna@gmail.com
⁵ Ph.D. Civil Engineering, Consulting Engineers – Brazil; e-mail: jelopes1@gmail.com

ABSTRACT

This paper presents the impact of the data quality from hydroelectric plants in the analysis of their past operation. A middle term simulation tool has been applied to Brazilian hydroelectric plants which are under the coordination of the Electric System National Operator (ONS). In order to analyze the impact of the data quality this simulator is used to reproduce the past operation of the plant, once with official data and average productivity and the next with adjusted data and variable overall efficiency. The results show that the use of both consolidated data and variable overall efficiency reduces the errors between recorded and simulated variables bringing the plant’s simulated and real operation closer.

Keywords: Hydroelectric Plants, Data Quality, Simulation Tool, Brazilian Model, Operation Planning.

1. INTRODUCTION

The power generated by a hydroelectric plant is a function of the water discharge by the turbines, the difference between the forebay and tailrace levels, the penstock head loss and the machines’ efficiency involved in the process. The level of accuracy used to represent the hydropower generation function depends on the problem addressed. For example, in the middle term planning, it is quite usual to consider an average productivity for the plant. While, in the short-term, usually a variable efficiency is considered for each machine.

This paper focuses on the middle term and its characteristics. The main objective is to compare the past operation analysis of a plant alternating: official data and consolidated data. The tool used to reproduce the past operation is a middle term simulator of the hydroelectric plants operation.

The first section presents the Brazilian model for middle term hydroelectric operation planning. The second details the concept and calculation of variable overall efficiency. The third describes the middle term simulator used in this paper. The fourth shows a graphic and numerical analysis on the impact of the data quality in the past operation reproduction. Conclusions are given in fifth section.

2. BRAZILIAN MODEL FOR MIDDLE TERM HYDROELECTRIC OPERATION PLANNING

The most common formulations of the Brazilian system for hydroelectric operation planning include: the production function and the water balance equation [12].

The goal of the production function is to quantify the power generation of a hydroelectric plant, as Eq. (1).

\[ p = k \cdot \eta \cdot [h_{fb}(x) - h_{fr}(u) - h_{pl}] \cdot q \]  

where:

- \( p \) is the instantaneous power obtained in the conversion process of the hydraulic...
potential energy to electric energy (MW).

\( k \)  
Is the gravity constant, multiplied by the water specific weight and divided by \( 10^6 \). Its value is 0.00981 (MW/(m³/s)/m).

\( x \)  
Is the water storage in the reservoir of the plant (hm³).

\( x_0 \)  
Is the forebay elevation which is function of the water storage \( x \) (m).

\( u \)  
Is the water release of the plant, that is, the sum of the water discharge by the turbines and the water spillage (m³/s).

\( h_{fb}(u) \)  
Is the tailrace elevation which is function of the water release \( u \) (m).

\( h_{pl} \)  
Is the penstock head loss which is function of the water discharge (m).

\( q \)  
Is the water discharge by the turbines of the powerhouse (m³/s).

The water balance equation, Eq. (2), is used to calculate the water mass conservation balance of the reservoir.

\[
x = x_0 + \left( y + \sum_{j \in \Omega} u_j - (q + s + ev + uc) \right) \cdot \frac{\Delta t}{10^6} \quad (2)
\]

where:

\( x_0 \)  
Is the reservoir volume at the beginning of period \( t \) (hm³).

\( y \)  
Is the incremental water inflow to reservoir during the period \( t \) (m³/s).

\( \Omega \)  
Is the upstream plants index set from the analyzed plant.

\( s \)  
Is the water spillage during the period \( t \) (m³/s).

\( ev \)  
Is the reservoir evaporation during the period \( t \) (m³/s).

\( uc \)  
Is the use of the reservoir’s water without the purpose of generating energy, such as: urban water supply, irrigation and navigation during the period \( t \) (m³/s).

\( \Delta t \)  
Is the size of the period \( t \) (s).

In order to aid the calculation of the parameters involved in the equations presented above are used seven physical functions: area-level polynomial, level-volume polynomial, level-release polynomial, maximum power function, maximum water discharge function, efficiency function and penstock head loss function [1].

\( \eta \)  
Is the constant or variable efficiency of the plant in the conversion process of the mechanical energy to electrical energy.

\( h_{pl}(x) \)

3. CONCEPT AND CALCULATION OF VARIABLE OVERALL EFFICIENCY

The overall efficiency includes the losses and the efficiencies involved in the operation. The use of the overall efficiency simplifies the production function, Eq. (1), without compromising the planning and the operation record of the plant.

Fig. 1 presents the simplification of the Eq. (1) using the concept of overall efficiency, \( \eta^G \).

\[
P = k \cdot \eta^G \cdot [h_{fb}(x) - h_{pl}(u)] \cdot q
\]

Fig. 1. Production function, Eq. (1), using the concept of overall efficiency.

The variable overall efficiency is represented as a hill curve matrix. This matrix can be a function of the gross head and the power output. For the attainment of the overall efficiency matrix using the data recorded by the plant can be used an optimization method, such as the “Solver” tool in Excel [3]. The objective function is to optimize the cells of the overall efficiency matrix in order to minimize the sum of squared error between the plant’s overall efficiency for the selected record and the overall efficiency calculated, as Eq. (3).

\[
\min \sum_{i=1}^{n} [\eta^G(i) - \eta^G(i)_{\text{calc}}]^2 \quad (3)
\]

where:

\( n \)  
Is the number of operations recorded in the plant’s database.

\( i \)  
Is the index of the operation recorded in the plant’s database.

\( \eta^G(i) \)  
Is the plant’s overall efficiency for the record of index \( i \).
\( \eta_{G(i)}^{\text{calc}} \) is the plant’s overall efficiency calculated using the overall efficiency matrix for the power output and the gross head of the record of index \( i \).

4. MIDDLE TERM SIMULATOR

The middle term simulator of the hydroelectric plants operation represents in detail the active operational restrictions on this horizon using weekly or monthly data. It can be used for planning the future operation or reproducing the past operation of a period. When it is used for reproducing the past operation it performs the function of a tool for data analysis like others mentioned in Hidalgo 2004 [5]; Hidalgo et al. 2009-A [6]; Hidalgo et al. 2009-B [7] and Hidalgo et al. 2009-C [8]. Its simulation process is based on the production function, Eq. (1), and the water balance equation, Eq. (2).

The software project and the computational implementation of this simulator use the Object-Oriented Paradigm [3], the C++ Programming Language [10] and the Structured Query Language (SQL) [4].

For studies of this paper, the simulator’s aim is to reproduce the water discharge trajectory from the initial volume, the trajectories of generation, water spillage and water inflow, as Fig. 2. The advantages of this type of application are: it shows the impact of the data inconsistency in the plant’s water balance and it can be used for planning the future operation.

5. GRAPHIC AND NUMERICAL ANALYSIS ON THE IMPACT OF THE DATA QUALITY IN THE PAST OPERATION REPRODUCTION

The middle term simulator was applied to a large Brazilian hydroelectric plant. The data recorded by the plant from 09/01/06 to 08/31/07 were compared to the resulting data from the simulation for the same period. The comparison was made in two situations. In the first the simulator worked with the official data provided by the company responsible for the operation. In the second the simulator used the consolidated data obtained according to the methodology presented in Hidalgo et al. 2009-D. Basically, the differences between the official and consolidated data are the six physical functions involved in the hydroelectric operation planning: area-level polynomial, level-volume polynomial, level-release polynomial, maximum power function, maximum water discharge function and efficiency function.

A - Simulation using official data and average productivity

In this simulation were used the official physical data of the company and the average productivity of the plant. The simulator’s goal is to reproduce the water discharge trajectory recorded by the plant, as Fig. 2.

Fig. 3, 4 and 5 show the recorded and simulated trajectories. It is possible to notice that the official physical data recorded by the plant are not coherent with the reality of the operation because the simulated trajectories moved away from the recorded trajectories.

Fig. 3. Comparison between the overall efficiency trajectories.

Fig. 4. Comparison between the water discharge trajectories.
The physical information that most influenced the result of this study was the plant’s productivity, Fig. 3. Since it is overestimated, the simulator saves reservoir’s water, Fig. 4, in order to produce the recorded generation. This justifies the increase of the forebay level, Fig. 5.

B - Simulation using consolidated data and variable overall efficiency

In this simulation were used the consolidated physical data of the plant and the variable overall efficiency calculated according to Eq. (3). Again, the simulator’s goal is to reproduce the water discharge trajectory recorded by the plant, as Fig. 2.

Fig. 6, 7 and 8 show the recorded and simulated trajectories of the plant.

These three figures show a strong coherence between the recorded data and their reproduction by simulation. This demonstrates the advantages of the use of variable overall efficiency even in the middle term.

In order to show a numerical analysis of the results presented above, Table I presents a statistical summary of the sum and mean squared error between the recorded and simulated variables when the simulator’s goal was water discharge reproduction.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Recorded Data x Simulated Data</th>
<th>Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Official</td>
<td>Adjusted</td>
</tr>
<tr>
<td>$h_{fb}(x)$</td>
<td>32.09</td>
<td>0.04</td>
</tr>
<tr>
<td>$h_{tr}(u)$</td>
<td>5.81</td>
<td>0.17</td>
</tr>
<tr>
<td>$q$</td>
<td>4,609,516.29</td>
<td>16,089.42</td>
</tr>
<tr>
<td>$s$</td>
<td>3,713,950.31</td>
<td>0.01</td>
</tr>
<tr>
<td>$\eta^G$</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>Mean squared error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h_{fb}(x)$</td>
<td>1.57</td>
<td>0.06</td>
</tr>
<tr>
<td>$h_{tr}(u)$</td>
<td>0.70</td>
<td>0.12</td>
</tr>
<tr>
<td>$q$</td>
<td>619.78</td>
<td>36.62</td>
</tr>
<tr>
<td>$s$</td>
<td>556.32</td>
<td>0.03</td>
</tr>
<tr>
<td>$\eta^G$</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>Mean reduction of the sum squared error</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The first column presents the analyzed variables. The second shows the sum and mean squared error between recorded data and simulated data using official information and plant’s average productivity. The third shows the sum and mean squared error between recorded data and simulated data using consolidated information and plant’s variable overall efficiency. The fourth column presents the error reduction between the second and third columns.

The numbers in Table I show that the data quality has great influence in the past operation
analysis of a plant. Some minor differences presented in the third column of the table are explained by accuracy of 0.1 MW used by the simulator. The differences related to the tailrace level and water discharge are explained by the fact that the plant’s tailrace is represented by a scattered cloud of points. It is believed that an accuracy technique for measuring of water discharge reduces the spaces among the points of the cloud further improving the results.

6. CONCLUSIONS

This paper compared the impact of the data quality in the past operation analysis of a large Brazilian hydroelectric plant. The analysis was made using a middle term simulator. The objective of this simulator was to reproduce the plant’s operation from 09/01/06 to 08/31/07 using monthly data. The simulator’s input data were classified in two kinds: official data using the plant’s average productivity and consolidated data using the plant’s variable overall efficiency.

The results were presented in the form of graphics and table. They all confirmed the importance of the data quality. For the analyzed study the mean reduction of the sum squared error between the recorded and simulated variables was of 99.16%.

The impact of the data quality in the past operation analysis indicates that others computational models used by the energy sector for optimization, simulation and streamflow forecasting may present dubious results due the quality of the data provided for them. Therefore, the search for the improvement of data is important to the choice of an economic and reliable operation policy for the hydroelectric system.

7. ACKNOWLEDGMENTS

The research reported herein was supported by the CNPq and CAPES, Brazilian Government agencies dedicated to the development of science and technology, which have funded, at different times, the Ph.D. studies of the first writer.

8. REFERENCES


