Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio

doi: 10.3850/978-981-18-8071-1_P717-cd

A new aggregation methodology for the HPP Health Index

Flávia Barbosa
INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal. E-mail: flavia.barbosa@inesctec.pt

Felipe Yamada
INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal. E-mail: felipe.yamada@inesctec.pt

Armando Leitão
INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal. E-mail: afleitao@fe.up.pt

Luís Guimarães
INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal. E-mail: luis.guimaraes@inesctec.pt

Hydroelectric Power Plant (HPP) supports European Union Electric Power System flexibility with various services (regulation capability, fast frequency control, fast start/stop, fast generating to pumping modes transition, high ramping rate, inertia emulation, and fault ride-through capacity, among others). New technology solutions, such as variable speed, are being studied to provide further flexibility in the framework of the XFLEX project. However, these additional capabilities impose new challenges on HPP's Operations and Maintenance (O&M). This work aims to increase the HPP's availability under this new paradigm. Proper health indexes (HI) should reflect the machine degradation, which is a critical component in health monitoring, fault diagnosis, and remaining useful life prediction. We explore indicators related specifically to Hydraulic Machines, namely the mechanical efficiency and the discharge of run-of-river Kaplan turbines, to develop a model to calculate the HPP global health index as a measurement of the impact in the HPP condition of each selected operating point for the energy production. This work proposes using Data Envelopment Analysis (DEA) to analyse the weights when aggregating the indicators into a global health indicator. Analyzing the global HI improves the equipment's performance, leading to operating points with better power production.

Keywords: Health Index, Hydroelectric Power Plant, Asset Management, Data Envelopment Analysis.

1. Introduction

Hydroelectric Power Plant (HPP) supports European Union Electric Power System (EPS) flexibility in regulation capability, fast frequency control, fast start/stop, fast generating to pumping modes transition, high ramping rate, inertia emulation, fault ride-through capacity, etc. New technology solutions, such as variable speed, are being studied to provide further flexibility. However, these additional capabilities pose new challenges to the Operations and Maintenance (O&M) of HPP.

This work aims at studying and optimising maintenance plans to decrease the outage time and increase the availability of the HPP under this new paradigm.

Proper HIs should be constructed to reflect the machine degradation, which is a critical component in health monitoring, fault diagnosis and remaining useful life (RUL) prediction. HIs can be classified into three categories: mechanical signal process-based, model-based and machine learning-based. This work considered the model-based HI, which consists of a database able to represent the behaviour of the main variables describing the HPP in the different operating points, numerical simulation results, and field tests and in an advanced control fed by the meta models designed to regulate the HPP units for reaching the "optimum operation point". By optimum operation, we mean the operation of one or multiple units that fulfill the given overall set-point while minimising a particular objective function. Therefore, the concept of optimum operation depends on the definition of the cost function. The cost function will be defined as a weighted sum of the degradation of the different HPP components, following Gerini et al. (2021). The aim is to extend the remaining lifetime by minimising this cost function. The methodology developed will give information on how to build the cost function and define the weights considering the various needs of the different HPPs.

This work focuses on exploring indicators related specifically to Hydraulic Machines, namely the mechanical efficiency and the discharge of run-of-river Kaplan turbines. The goal is to develop new models to calculate the HPP health index (HI) as an indicator of the impact in the HPP condition of the selected operating points. These models are then to be incorporated into the advanced control optimisation algorithm. Specific KPIs evaluate the degradation process and are included in the decision process to define the best maintenance policy and a risk analysis. An overall weighted health index, aggregating those KPIs, may indicate the equipment's condition. This work proposes using Data Envelopment Analysis (DEA) to analysing the weights for the KPIs. Data Envelopment Analysis is a linear programming method introduced by Charnes et al. (1978) for efficiency analysis. It is a non-parametric frontier technique that does not require a specific functional form of a production function to measure efficiency. Instead, it evaluates the relative efficiency of the production units under analysis. When it comes to selecting weights, if using a standard fixed weighting scheme, they must be determined prior to running the model, typically through expert input. This is frequently criticized due to the inherent subjectivity involved in defining them. There is also frequently a lack of consensus among experts regarding the appropriate weights to be used in the aggregation functions. Furthermore, these methods are incapable of taking unit-specific characteristics into account. As a result, the importance level assigned to each indicator by each unit is ignored, complicating the investigation of root causes of poor performance. In the absence of reliable and consensual information about the weights, the DEA model endogenously selects the weights that maximise the HI score for the entity under assessment. Thus, each unit can be assessed with its weights, emphasising good performance indicators. It is expected that the defined HI can diagnose and prognose the equipment's health index, leading to operating points that consider the trade-off system flexibility and power production with the system's reliability.

2. Condition Monitoring in Hydroelectric Power Plants

The initial investment costs of an HPP are relatively high; however, HPPs have a very long lifespan. Initially, operations are stable and problem-free, and the number of faults is low. However, due to the degradation of the system's elements, the number of faults increases with time. In this context, Selak et al. (2014) presented a condition monitoring and fault diagnostics system for hydropower plants by comparing data recorded during fault-free operation and the operational phase.

The mechanical component is one of the core parts of the hydropower plant, stressing the importance of hydro-condition monitoring. When running in a low-efficiency district or low-head district, some turbines cause vibrations and air corrosion damage. Therefore, the awareness of the turbine situation plays an important role. Unreasonable design, manufacturing defect, or installation defect may cause problems in the turbine operation. Zhang and Tongji (2011) redesigned the hardware of the hydropower station hydraulic monitoring system Tianqiao hydropower to achieve multi-channel data acquisition, multi-functional analysis, and visualisation operations. To minimize the runner damage, the penstock fatigue and the water losses in a pump-turbine, Schmid et al. (2022) proposed an optimisation approach to determine a start-up sequence. Valentín et al. (2022) compared the hybrid mode and the standard mode (non-hybrid) in terms of mileage and wear and tear of the guide vanes and runner blades servomotors. They have observed that with a battery in parallel, the units can provide this service without regulating that much the power since this task is done by the battery. This hypothetically reduces the number of manoeuvres in the regulation systems and therefore reducing their wear and tear.

The health index is based on expert knowledge or data-driven models of specific transformer subsystems and combines multiple information sources and generate a consistent health state indicator for asset management. Aizpurua et al. (2019) considered different sources of uncertainty in power transformers and inferred a HI value with confidence intervals for decision-making under uncertainty.

3. Methodology

The methodology of this paper has 2 steps. First, we use Data Envelopment Analysis to define weights for the HI. Next, theses weights are used in a control model to optimise the Frequency Containment Reserve (FCR) provision of double-regulated turbines, typical machines deployed in low head hydropower plants.

3.1. *Data Envelopment Analysis*

Data envelopment analysis (DEA) is a data-oriented approach that allows to evaluate the performance of a set of entities. It is done by estimating the best practice frontier shared by production units consuming a given amount of inputs and transforming them into a certain amount of outputs. The distance to such best practice frontier is a measure of efficiency, and it reflects how much inputs can decrease while maintaining the same production of outputs, or by how much outputs can be increased without the need to consume extra resources.

Consider a set of entities $J = \{1, 2, \ldots, n\}$, each consuming different amounts of m inputs to produce s outputs.

Let $I = \{1, 2, \ldots, m\}$ be the set of inputs and $R = \{1, 2, \ldots, s\}$ the set of outputs. v_i and u_r are the weights given, respectively, to inputs $i \in I$ and outputs $r \in R$. x_{ij} $(\in \Re_0^+)$ are the observed inputs i and y_{rj} $(\in \Re_0^+)$ are the observed outputs r of entity $j \in J$. To evaluate the efficiency of each entity j_0, the ratio of weighted outputs to weighted inputs of the entity under assessment j_0 must be maximised, and the similar ratios defined for all other entities must be less than or equal to unity. Let x_{i0} and y_{j0} be the inputs i and the outputs j of entity 0 under assessment. The DEA ratio model is formulated as follows:

$$
\begin{aligned}
\max \quad & \frac{\sum_{r=1}^{s} u_r y_{rj_0}}{\sum_{i=1}^{m} v_i x_{ij_0}} \\
\text{s.t.} \quad & \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \ \forall j \in J \\
& u_r \geq 0 \ \forall r \in R \\
& v_i \geq 0 \ \forall i \in I
\end{aligned} \tag{1}
$$

Model (1) maximizes its efficiency for each unit j_0. The decision variables are the weights associated with each input and output. The maximum achievable efficiency is 1, indicating that the weights assigned to a DMU cannot yield an efficiency value greater than 1 when applied to all the DMUs (the first constraint). The weights must be greater than or equal to 0 (second and third constraints). The optimal weights defined in the model aggregate multiple outputs into a virtual output and multiple inputs into a virtual input. The ratio of the virtual output to the virtual input represents a single efficiency measure. The optimal weights are flexible and vary for each DMU, highlighting their best possible performance. Model (1) is not linear, but it can be easily corrected. When converting the model to linear programming, the efficiency of entity j_0 can be determined either with an input or output orientation. The choice of orientation depends on the characteristics of the problem under study. In the input-oriented formulation, the efficiency is given by the minimal factor by which all inputs of entity j_0 can be proportionally decreased without decreasing any output level. In the output-oriented formulation, the efficiency is given by the inverse of the maximum factor by which all outputs of entity j_0 can be proportionally increased without increasing any input level.

3.2. *Vogelgrün Case Study*

The proposed case study of this report is the Vogelgrün run of river (RoR) hydropower plant. Built in 1959 on the Grand Canal d'Alsace by the Rhine river, Vogelgrün RoR features four low head Kaplan units. It also includes two locks of major importance for international river navigation with more than 20000 boats yearly. The units have long been used for flow and water level control, whereas new needs for grid support are emerging.

The methodology proposed follows three steps.

In the first step, we generate a database of the relevant state variables, representing the operational parameters of the hydroelectric units (H, Q^{SET}, α, and β).

database variable	minimum	maximum
H	10	12
Q^{SET}	300	350
α	60%	100%
β	60%	100%

To have a comprehensive knowledge of the hydropower plant operation, To have a comprehensive knowledge of the hydropower plant operation, Vagnoni et al. (2021) developed metamodels using Multivariate Adaptive Regression Spline (MARS) for estimating mechanical efficiency η (equation 2) and discharge Q (equation 3) depending on the recorded H, Q^{SET}, guide vanes angle α and blades angle β. The MARS modeling allows for evaluating the influence of each independent variable by using only an initial exploration data set which is well suited to validate the independent variables selected in the study.

$$\eta = f_1(H, Q^{SET}, \alpha, \beta) \tag{2}$$

$$Q = f_2(H, Q^{SET}, \alpha, \beta) \tag{3}$$

We use the previously obtained state variables data to fed the metamodels to provide the database variables of mechanical efficiency and then compute (HI_η) and discharge (HI_Q).

We also calculate the power generation P of a hydroelectric unit is computed as in (4):

$$P = \rho \times g \times H \times Q \times \eta \times \eta_e \tag{4}$$

Where the density of the water ρ is considered equal to 1000 kg/m^3 and the gravitation acceleration 9.81 m/s^2. η_e represents the efficiency of the synchronous generator. The net head H is measured onsite and collected in the operational statistics, while the discharge Q and the hydraulic machine efficiency η are estimated by the metamodels

Next, Gerini et al. (2021) developed a control strategy for optimal asset management of hydroelectric units in run-of-river hydropower plants. The control-oriented modelling methodology integrates the operational parameters of the hydroelectric unit in an optimisation algorithm steering the advanced control of the units.

Model (4) optimise the Frequency Containment Reserve (FCR) provision of double-regulated turbines, typical machines deployed in low head hydropower plants. They define the control problem as finding the combination of guide vanes α and blades opening angles β that maximises the efficiency and the discharge tracking, for a given discharge set-point Q^{SET} and external condition n_{ED} (i.e. head H and rotational speed n).

$$\begin{aligned}
\text{Min}_{\alpha,\beta} \quad & \omega_\eta[1 - \eta^*(\alpha, \beta, n_{ED,s})] + \\
& \omega_Q[Q_s^{SET} - \bar{Q}^*(\alpha, \beta, n_{ED,s})|_{\alpha_{s-1},\beta_{s-1}}]^2 \\
\text{s.t.} \quad & \alpha_{s-1} + \nu_{\alpha c}\Delta t \leq \alpha \leq \alpha_{s-1} + \nu_{\alpha o}\Delta t \\
& \beta_{s-1} + \nu_{\beta c}\Delta t \leq \beta \leq \beta_{s-1} + \nu_{\beta o}\Delta t \\
& \alpha, \beta \in \Omega_{\alpha,\beta}
\end{aligned} \tag{5}$$

where Δ = 1s, and $\nu_{\alpha o}$, $\nu_{\alpha c}$ are respectively the normalised maximum speeds in opening and closing of the servomotor acting on the guide vanes, and $\nu_{\beta o}$, $\nu_{\beta c}$ are the corresponding quantities for the blades servomotor. The operation must be within feasible positions of guide vanes and blades, and speed of the servomotors is limited, to avoid changes in the moving organs which are physically impossible. Gerini et al. (2021) did not discussed the choice of the weights ω_η and ω_Q. Therefore, we focused on measuring the impact of degradation on the power generation through the weights analysis. In this case study, the degradation is represented by the health indexes for mechanical efficiency $HI_\eta = 1 - \eta^*(\alpha, \beta, n_{ED,s})$ and discharge $HI_Q = [Q_s^{SET} - \bar{Q}^*(\alpha, \beta, n_{ED,s})|\alpha_{s-1}, \beta_{s-1}]^2$

In the second step, each triple (HI_η, HI_Q,P) is defined as a entity, so we can run the DEA model to obtain values for the weights ω_η, ω_Q.

We consider as inputs the degradation (HI_η and HI_Q) and as output the power generated (P), as shown in Figure 1:

Fig. 1. Input and output configuration

To minimise the inputs levels (degradation) with an assumption of fixed outputs levels (power generation), we use the DEA linear programming model with an input orientation, as formulated in (6):

$$\begin{aligned} \max \quad & \sum_{r=1}^{s} u_r y_{rj_0} && (6) \\ \text{s.t.} \quad & \sum_{i=1}^{m} v_i x_{ij_0} = 1 \\ & \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \ \forall j \in J \\ & u_r \geq 0 \ \forall r \in R \\ & v_i \geq 0 \ \forall i \in I \end{aligned}$$

It is necessary to make sure that there is not much imbalance in the data sets, i.é., the data is at the same or similar magnitude. They suggest normalising the data by the mean to tackle this problem (find the mean of the data set for each input and output and divide each input or output by the mean for that specific factor).

Finally, in the third step analyse the weights to measure the trade-off between input consumption (HI_η and HI_Q) and output production (P).

For each entity j_0, the efficiency score is computed following Model (6) as:

$$eff = \frac{u_P \times P_{j_0}}{v_\eta \times \eta_{j_0} + v_Q \times Q_{j_0}} \tag{7}$$

To measure how much degradation (in terms of mechanical efficiency and discharge) we are generating for each unit of power produced, we can compute the ratios defined respectively in equations (8) and (9):

$$ratio_\eta = \frac{v_\eta}{u_P} \tag{8}$$

$$ratio_Q = \frac{v_Q}{u_P} \tag{9}$$

We then run the control model considering the correct proportion for the weights, in order to improve the power generation.

4. Results

We generated 1019 combinations of α and β. We use the metamodels to obtain η and Q and to compute:

- mechanical efficiency: $HI_\eta = 1 - \eta$
- discharge: $HI_Q = (Q - Q^{SET})^2$
- power: $P = \rho \times g \times H \times Q \times \eta$

We summarise such information in Table 1.

Table 1. Dataset generation

	α	β	Q	HI$_\eta$	HI$_Q$	power
mean	75.1	77.4	300.0	43.4	4242.3	22684.5
std dev	11.9	12.5	0.0	42.8	3357.6	3022.6
min	60.0	60.0	300.0	0.0	0.0	16477.5
25-perc	66.0	66.0	300.0	10.5	1459.9	20325.5
75-perc	82.0	89.0	300.0	62.8	6445.9	24736.0
max	100.0	100.0	300.0	212.7	14618.4	31211.3

Next, we mean normalised the data to make sure the data is of similar magnitude across, as we show in Table 2.

Table 2. HI computation

	HI$_\eta$	HI$_Q$	power	HI$_\eta$ norm	HI$_Q$ norm	power norm
mean	43.4	4242.3	22684.5	1.0	1.0	1.0
std dev	42.8	3357.6	3022.6	1.0	0.8	0.1
min	0.0	0.0	16477.5	0.0	0.0	0.7
25-perc	10.5	1459.9	20325.5	0.2	0.3	0.9
75-perc	62.8	6445.9	24736.0	1.4	1.5	1.1
max	212.7	14618.4	31211.3	4.9	3.4	1.4

In Table 3 we show the results obtained from the DEA model, and compute the rations as in equations (8) and (9).

Table 3. Weights summary

Eff	w(eta)	w(Q)	w(power)	racio_{eta}	racio_Q	
mean	0.0	7.9	949.8	0.0	12.7	1.1
std dev	0.1	102.6	30101.2	0.1	139.8	1.0
min	0.0	0.0	0.0	0.0	0.2	0.0
25-perc	0.0	0.1	0.2	0.0	0.6	0.3
75-perc	0.0	2.7	2.1	0.0	4.4	1.7
max	1.0	3085.2	961363.0	0.8	4142.1	4.7

We notice that in the most times, the ratio (9) is 21 times greater than the ratio (8). From that, we suggest the weights $w_\eta = 1$ and $w_Q = 21$. That is, the objective function in control model (4) is:

$$\text{Min}_{\alpha,\beta} \;\; [1 - \eta^*(\alpha, \beta, n_{ED,s})] + 21[Q_s^{SET} - \bar{Q}^*(\alpha, \beta, n_{ED,s})|\alpha_{s-1}, \beta_{s-1}]^2 \tag{10}$$

We compare the power generated wen running for one day (01-jan-2018) with $w_\eta = w_Q = 1$ and 10.

For $w_\eta = w_Q = 1$, the total power generated in this day was 315956693.6W; for $w_\eta = 1$ and $w_Q = 21$, the total power generated in this day was 315956711.9W. Therefore, with the new weights we manage to improve power generation in 0.00001%.

Fig. 2. HI_η comparison

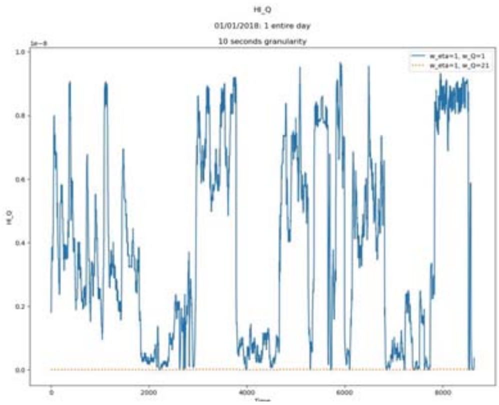

Fig. 3. HI_Q comparison

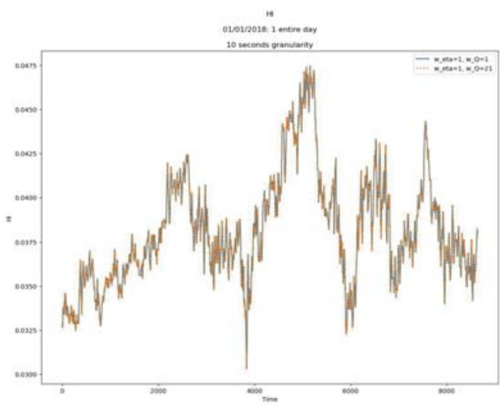

Fig. 4. Global HI comparison

The increase of frequency control actions enlarge the discrepancy between discharge set-point usually established by day-ahead markets and real value of the discharge. Although the improvement in the power generation may seem small, it was possible to promote higher stability in the discharge, as we show in Figure 3. A considerable deviate of the discharge from its expected value because of FCR provision, could cause the alteration of the river head as well. For the mechanical efficiency (Figure 2) and the global health index (Figure 4) it is not possible to note a significant difference in the stability over time.

5. Conclusion

This work proposed the use of Data Envelopment Analysis (DEA) to analyse the weights when aggregating the indicators into a global health indicator. The wear and tear were included in the health index function to consider the minimization of the damage function in the choice of the set-points (values for angles α and β) given by the optimization algorithm. As a result, the operational parameters characteristics of the hydroelectric unit operation can be successfully integrated into an advanced control based and to define the best operating set-point of the unit. Analyzing the global HI improves the equipment's performance, leading to operating points with better power production.

6. Future work

In the following steps, we intend to extend and run the tests for data for the other months of the year and compare them with cost functions dependent on the amount of power generated. It is noteworthy that the proposed methodology can be extended to handle more than two inputs. That means we can incorporate other HI measures, such as damage and the number of starts and stops, which is the case of the Frades 2 HPP in Portugal.

Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857832.

References

Aizpurua, J. I., B. G. Stewart, S. D. McArthur, B. Lambert, J. G. Cross, and V. M. Catterson (2019). Improved power transformer condition monitoring under uncertainty through soft computing and probabilistic health index. *Applied Soft Computing 85*, 105530.

Charnes, A., W. W. Cooper, and E. Rhodes (1978). Measuring efficiency of decision making units. *European Journal of Operational Research 2*, 429–444.

Gerini, F., E. Vagnoni, R. Cherkaoui, and M. Paolone (2021). Improving frequency containment reserve provision in run-of-river hydropower plants. *Sustainable Energy, Grids and Networks 28*, 100538.

Schmid, J., S. Alligné, D. Biner, C. Münch-Alligné, N. Hugo, and C. Nicolet (2022). Optimization of turbine start-up sequence of a full size frequency converter variable speed pump-turbine. In *IOP Conference Series: Earth and Environmental Science*, Volume 1079, pp. 012109. IOP Publishing.

Selak, L., P. Butala, and A. Sluga (2014). Condition monitoring and fault diagnostics for hydropower plants. *Computers in Industry 65*(6), 924–936.

Vagnoni, E., F. Gerini, R. Cherkaoui, and M. Paolone (2021). Digitalization in hydropower generation: development and numerical validation of a model-based smart power plant supervisor. In *IOP Conference Series: Earth and Environmental Science*, Volume 774, pp. 012107. IOP Publishing.

Valentín, D., A. Presas, M. Egusquiza, and C. Valero (2022). Hybridization in kaplan turbines. wear and tear assessment. In *IOP Conference Series: Earth and Environmental Science*, Volume 1079, pp. 012108. IOP Publishing.

Zhang, W. and Z. Tongji (2011). Development and application of hydropower hydro monitoring system. *Procedia Engineering 15*, 807–811.