15th Triennial World Congress, Barcelona, Spain

THE INFLUENCE OF HYDROLOGIC INFORMATION IN LONG-TERM HYDROTHERMAL SCHEDULING

L. Martinez [*]; **R. Ballini** [**]; **S. Soares** [*]

[*] *Department of Engineering System*
School of Electrical and Computer Engineering
State University of Campinas

[**] *Department of Computer Engineering and Industrial Automation*
School of Electrical and Computer Engineering
State University of Campinas
e-mail: {martinez, ballini, dino}@densis.fee.unicamp.br

Abstract: This paper presents an evaluation of the benefits of streamflows forecasting in long-term hydroelectric scheduling problem. In the approach considered, at each stage of planning a forecast of the future inflows is made and an operational decision for the following stage is obtained by a deterministic optimization model, in a partial open-loop feedback control framework. The influence of the forecasting model in the performance of this control policy is analysed by simulation using historical inflows record. The effectiveness of the approach was measured using the mean and standard deviation values for hydro generation and operational costs during the planning period, taking into account the hydroelectric plants of the Southeast Brazilian System as a case study. *Copyright © 2002 IFAC.*

Keywords: Hydrothermal Power Systems; Power Generation; Prediction Methods; Nonlinear Programming.

1. INTRODUCTION

Long term hydrothermal scheduling (LTHS) is a complex problem due to various aspects involved, including the randomness of inflows into the hydro plants, the interconnection of hydro plants located in a cascade and the nonlinearity of hydro production and thermal cost functions. It is important that these aspects be regarded simultaneously for an adequate modeling of the problem.

In this paper, the approach considered in the solution of the LTHS problem combines a deterministic optimization of the problem (Hanscom *et al.*, 1980), (Bissonnette *et al.*, 1986), (Oliveira and Soares, 1990) with inflows furnished by a forecasting model. The deterministic optimization model permits the representation of the hydro system in detail, considering each hydro plant individually, including its operational constraints and nonlinear production characteristics. Moreover, the stochastic model contemplated for the representation of the inflows can be quite general, based on any methodology and specific for each hydro plant in the system.

The optimal decision based on the current forecast of future inflows is implemented until a new inflow forecast becomes available based on the latest available information in the system. In this way, the operational policy is determined within the framework of partial open-loop feedback control (Bertsekas, 1995), (Martinez and Soares, 2002).

The goal of the present paper is to evaluate the benefits of streamflow forecasting in LTHS, taking into account the partial open-loop feedback control framework. Three different inflow forecasting models are considered: the long-term average inflows, a lag-one periodic autoregressive model and a neurofuzzy network model.

The performance of the approach was evaluated by simulation using historical inflow records. Deterministic optimization, assuming perfect foresight of inflows during the planning period, was also considered in order to evaluate the influence of the randomness of inflows in the LTHS problem. A test system composed of hydroelectric plants located in the Southeastern Brazilian Power System was considered as a case study.

The paper is structured as follows: Section 2 describes the formulation of the LTHS problem. Section 3 presents the partial open-loop feedback control policy adopted. Section 4 presents the three different models for the stochastic representation of inflows. Section 5 shows the numerical results for the test system and finally Section 6 presents the conclusions of the study.

2. LTHS FORMULATION

The deterministic version of the LTHS can be formulated as the following nonlinear programming problem:

$$\min \sum_{t=1}^{T} \psi_t(G_t) + V(x_T) \qquad (1)$$

subject to:

$$G_t + P_t = D_t, \quad \forall t \qquad (2)$$

$$P_t = \sum_{i=1}^{I} p_{i,t}, \quad \forall t \qquad (3)$$

$$p_{i,t} = k_i h_{i,t} q_{i,t}, \quad \forall i, t \qquad (4)$$

$$h_{i,t} = \phi(x_{i,t}) - \theta(u_{i,t}), \quad \forall i, t \qquad (5)$$

$$x_{i,t} = x_{i,t-1} + y_{i,t} + \sum_{k \in \Omega_i} u_{k,t} - u_{i,t}, \quad \forall i, t \qquad (6)$$

$$u_{i,t} = q_{i,t} + s_{i,t}, \quad \forall i, t \qquad (7)$$

$$\underline{x_{i,t}} \le x_{i,t} \le \overline{x_{i,t}}, \quad \forall i, t \qquad (8)$$

$$\underline{u_{i,t}} \le u_{i,t} \le \overline{u_{i,t}}, \quad \forall i, t \qquad (9)$$

$$\underline{q_{i,t}} \le q_{i,t} \le \overline{q_{i,t}}, \quad \forall i, t \qquad (10)$$

$$s_{i,t} \ge 0, \quad \forall i, t \qquad (11)$$

$$x_{i,0} \ \text{given}, \quad \forall i, t \qquad (12)$$

The objective function (1) is composed of two terms which represent the operational cost during the planning period and the future cost associated with the final storage in the reservoirs.

The operational cost ψ_t represents the minimum cost of complementary non-hydro sources such as thermoelectric generation, imports from neighboring systems, and load shortage. The function ψ_t is determined by the economic dispatch of such sources and, as consequence, is a convex increasing function of the total non-hydro generation G_t.

The function $V(x_T)$ is a terminal condition which represents the future operational cost as a function of the final reservoir storage. This term is essential for the equilibrium between the use of water during the planning period and its use afterwards.

Equation (2) represents for each stage the power balance, where P_t is the total hydro geneeration and D_t is the total load demand.

Hydro generation at plant i and stage t is a nonlinear function represented by equation (4), where the constant k_i is the product of water density, gravity acceleration and average turbine/generator efficiency, $h_{l_{i,t}}$ is the water head function represented by equation (**??**) and $q_{i,t}$ is the water discharge through the turbines of the hydro plant i.

The water head is a function of the water storage in the reservoir, $x_{i,t}$ and water release from the reservoir $u_{i,t}$, where $s_{i,t}$ represents the water spillage from the reservoir. The forebay $\phi(.)$ and tailrace $\theta(.)$ elevations are fitted by polinomial functions.

The equality constraint in (6) represents the water balance in the reservoir at each stage t, where y_t is the incremental water inflow. Other terms, such as evaporation and infiltration have not been considered for the sake of simplicity.

Lower and upper bounds on variables, expressed by constraints (8)-(11), are imposed by physical operational constraints of the hydro plant, as well as other constraints associated with the multiple uses of water, such as irrigation, navigation and flood control.

3. PARTIAL OPEN-LOOP FEEDBACK CONTROL

In the partial Open-Loop Feedback Control (OLFC) approach (Martinez and Soares, 2002), the randomness of inflows is considered in an implicit way since stochastic variables are assigned to their expected values, provided by inflow forecasting models.

The problem is solved by a deterministic optimization model, and the optimal decision variable associated with the first stage is implemented. In order, to avoid error propagation, the scheme is repeated at each stage throughout the planning period.

The solution of the deterministic optimization model is obtained by a nonlinear network flow algorithm specially developed for hydrothermal scheduling (Oliveira and Soares, 1990).

One important issue in the design of the OLFC approach is the terminal condition $V(x_T)$, which establishes a trade-off between the benefits associated with the use of water for hydro generation during the planning period and the expectation of future benefits deriving from storage at the end of the planning period, both measured in terms of non-hydro generation savings.

One way of obtaining a proper terminal condition is to extend the end of the optimization period so that the influence of $V(x_T)$ on the decision during the first stage becomes negligible. However, this can be rather inconvenient since the extension of the planning period increases the forecasting errors. On the other hand, establishing a shorter planning period so that the forecasting model will be able to provide better performance would require an accurate estimation of the expected future operational cost since, in this case, the influence of the terminal condition on the decision of the first stage is crucial.

The terminal condition considered in this paper will try to maintain the storage of the hydro plant as full as possible at the beginning of the dry season, which is May for the hydro plants considered in the case study. This terminal condition is inspired on the solution of the deterministic optimization model over the historical inflow records.

Assuming that T represents the next month of April, the solution of the deterministic optimization model is obtained considering the terminal condition $V(x_T) = M(\overline{x} - x_T)$, where M is a positive constant large enough to ensure that the terminal condition prevails over the remaining objective function.

4. STREAMFLOW FORECASTING MODELS

The stochastic models considered in the OLFC policy are based on historical inflows. The monthly series of natural inflows are seasonal series and present a periodic behavior.

The first model adopted for the representation of the inflows is the Long Term Average (LTA) inflow, which corresponds to a very simple forecasting model. The intention is to design the system operation assuming that future hydrological conditions will be identical to the historical average values. Others forecasting models considered in the paper are a lag-one periodic autoregressive model and a neurofuzzy network model, which are briefly described.

4.1 Periodic Auto Regressive

Inflow series have been frequently represented by autoregressive models where the parameters have a periodic behavior, such as Periodic Auto Regressive models (PAR).

In this paper, a lag-one periodic autoregressive PAR(1) model was applied to the actual historical inflow records after normalization of the series by subtracting the expected value and dividing by the standard deviation. As a result, the PAR(1) model is represented by:

$$z_{t(r,m)} = \phi_m z_{t(r,m)-1} + a_{t(r,m)} \tag{13}$$

where,

ϕ_m is the autocorrelation coefficient of the normalized serie;

$a_{t(r,m)}$ is a sequence of uncorrelated random variables with distribution $N(0, \tau_m^{-1})$;

$z_{t(r,m)}$ represents the padronized serie

$$z_{t(r,m)} = \frac{y_{t(r,m)} - \mu_m}{\sigma_m} \tag{14}$$

where,

$y_{t(r,m)}$ is the inflow at time $t(r,m) = 12(r-1) + m$, with r being the year and m the month;

μ_m is the expected value of the inflow for month m;

σ_m is the standard deviation of the inflow for month m.

The autocorrelation coefficients ϕ_m were estimated using the Maximum Likelihood Estimate method as suggested in (Box *et al.*, 1994), (Vecchia, 1985).

4.2 Neurofuzzy Network Model

Another inflow forecasting model considered in this paper is based on a new class of Neurofuzzy Network (NN) proposed by (Figueiredo *et al.*, 1995). In these networks the essential parameters for modelling a fuzzy system, such as fuzzy rules and membership functions, are learned through a constructive learning method where neurons groups compete when the network receive a new input.

The NN has a feedforward architecture with five layers. The topology of the neurofuzzy network presents two essential features: the mapping of fuzzy rules into or from the network structure is direct, and the fuzzy inference and the neural processing are in complete agreement. Therefore, the approach proposed has a dual nature, i.e., it can be seen either as a neural fuzzy network or a fuzzy rule based system.

Fuzzy inference is a mapping from an observed nonfuzzy space $Z \subseteq \Re^n$ to the fuzzy sets in Z. A fuzzy set defined in Z is characterized by a membership function $F : Z \to [0,1]$ (Zadeh, 1978). The network emulates fuzzy reasoning mechanisms, encoding the fuzzy rule base in the form of "If a set of conditions is satisfied, Then a set of consequences is inferred".

The grade of membership of x_k in Z is z_k, i.e., $Z(x_k) = z_k$, if $x_k \in I_k = (x_I, x_F)$, where Z is a fuzzy set and x_k is a numerical value of the input space. The numerical value of the output y is determined by a sequence of stages, as proposed by (Pedrycz and Gomide, 1998), (Yager and Filev, 1994):

1. *Matching:* For each rule i and each antecedent j, compute the possibility measure P_j^i for fuzzy sets A_j and A_j^i, given by:

$$\mathrm{P}_j^i(\mathbf{x}) = \underset{k}{\mathrm{S}} \left\{ \mathrm{T} \left(A_j(\mathbf{x}), \mathrm{A}_j^i(\mathbf{x}) \right) \right\} \qquad (15)$$

where S is taken over all k, and $\mathbf{x} = (x_1, \ldots, x_M)'$ is the input vector, and T and S denote the t-norm and co-norm, respectively.

2. *Antecedent Aggregation:* For each rule i compute its activation level H^i, defined by:

$$\mathrm{H}^i(\mathbf{x}) = \underset{j}{\mathrm{T}} \{ \mathrm{P}_j^i(\mathbf{x}) \} \qquad (16)$$

3. *Rule Aggregation:* The output y is computed by:

$$y(\mathbf{x}) = \sum_{i=1}^{N} \mathrm{H}^i(\mathbf{x})\, w_i / \sum_{i=1}^{N} \mathrm{H}^i(\mathbf{x}) \qquad (17)$$

where w_i denotes the weight of the connection i.

The learning algorithm consists of the presentation of input/desired output pairs to the network. The t-th pair presented to the network is given by $(\mathbf{x}(\mathrm{t}),\ y_d(\mathrm{t}))$, $\mathbf{x}(\mathrm{t}) = (x_1(\mathrm{t}),\ \ldots,\ x_M(\mathrm{t}))'$ the input vector and $y_d(\mathrm{t})$ the desired output. The structure of the network changes during the training, which means that the number of fuzzy rules is not constant ($N = N(\mathrm{t})$, where t is the iteration index). When the network receives a new input the neuron groups compete. If one of the neurons group wins, the desired performance is satisfied and the parameters are adjusted. If the desired performance cannot be satisfied or if no neuron group matches the input a new group is added to the structure of the network.

The learning method acquires new knowledge whenever necessary. The strategie provides an automatic way to learn parameters for a fuzzy model without interference or participation of an expert.

This approach was applied to seasonal streamflow forecasting in (Ballini *et al.*, 2001) using a database of average monthly inflows of Brazilian hydroelectric plants. The results obtained with the model were compared to a multilayer feedforward network and periodic autoregressive models, showing a significant decrease in the forecasting errors.

5. SIMULATION RESULTS

In this work, the test system comprises the hydroelectric plants of the Grand River cascade, located in the Southeastern region of Brazil (Figure 1). The main characteristics of the hydro plants are presented in Table 1.

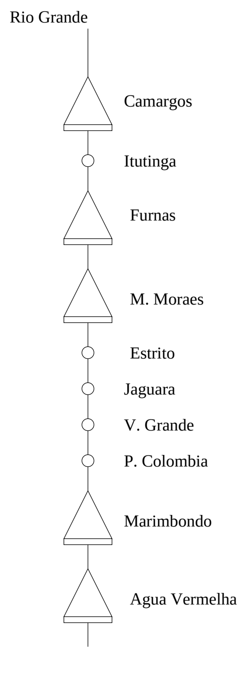

Fig. 1. Grande River cascade.

The operational cost ψ_t is obtained by an economic dispatch of the non-hydro sources available, resulting in a convex increasing operational cost function. For the Brazilian Southeastern Power System, an estimative of the operational cost is given by the following quadractic function:

$$\psi_t = 0.02\,(D_t - P_t)^2 \qquad (18)$$

The performance of the OLFC, considering different forecasting models, was analysed using a simulation model which reproduces the behavior of the

Table 1. Hydroelectric plant characteristic.

Name	Installed Capacity (MW)	Storage Capacity (hm^3)	Discharge max/min (hm^3/s)
Camargos	48	572	32/288
Itutinga	52	-	32/244
Furnas	1312	17217	196/1692
M. Moraes	478	2500	247/1328
Estreito	1104	-	252/2028
Jaguara	424	-	255/1564
Volta Grande	380	-	275/1584
Porto Colômbia	328	-	307/1988
Marimbondo	1488	5260	441/2944
Água Vermelha	1380	5169	510/2922

hydrothermal system. The simulations were made over the historical inflow records in the period from May 1931 to April 2000. A simulation with perfect forsight of the future inflows, named (PF), was also considered. The statistics of interest are the mean and standard deviation values of hydro generation and operational cost.

In a first study, the operation of a single hydro plant (Furnas) system and a constant load demand D_t of 1312 MW was considered. Table 2 presents the mean and standard deviation values of hydro generation and operational cost for this case.

Table 2. Numerical results for Furnas hydro plant.

Inflow	Hydro Generation		Op. Cost $[\times 10^6]$	
Information	Mean	St. Dev.	Mean	St. Dev.
PF	743.02	173.97	3.7725	1.9137
LTA	722.62	237.44	4.3023	2.8825
PAR(1)	725.41	248.91	4.3253	3.0426
NN	740.78	158.67	3.9880	2.3960

The results show higher average hydro generation with the use of the NN model. The standard deviation obtained in this case was smaller, and since the operational cost is convex and increasing, such flat behavior lead to smaller final operating costs. It is quite interesting to observe that the solution based on the NN model did not differ significantly from the solution obtained with perfect foresight of inflows.

Figure 2 shows the trajectories of water storage in the reservoir obtained by the OLFC policy with different forecasting models during the period from May 1950 to April 1960, which includes the critical inflow period of the Grand river.

Note in Figure 2 that from 1950 to 1952 and from 1959 to 1960, which corresponds to average periods, the differences among the forecasting models are relatively low, in contrast to the large diffences verified from 1952 to 1956, which corresponds to the driest period of the historical records. The OLFC policy associated with the NN model resulted in a solution that lead to higher storage levels in the reservoir, increasing the productivity of the plant and therefore its efficiency. The improved performance during the critical period, where the operational cost and the shortage risk are greater, makes the OLFC policy with NN forecasting model an adequate approach for LTHS.

The results also revealed that the operational policy based on the LTA forecasting model did not differ significantly from those obtained with the PAR(1) model. This suggests that the multiple steps ahead forecasting with the PAR(1) model quickly tend to the average inflow values.

In a second study, the system comprising the whole cascade and a constant load demand of 7188 MW is considered. The numerical results are presented in Table 3.

Table 3. Statistics of simulations for Grand River cascade.

Inflow	Hydro Generation		Op. Cost $[\times 10^6]$	
Information	Mean	St. Dev.	Mean	St. Dev.
PF	4858.7	875.03	6.5975	3.9153
LTA	4667.0	1157.40	8.1990	7.2450
PAR(1)	4670.7	1219.40	8.3365	7.6964
NN	4726.5	1166.50	7.9060	7.0584

In this case, the reduction observed in cost with the NN model were about 5% in relation to the PAR(1) model, showing a significant improvment due to the use of more efficient forecasting models in the OLFC policy.

6. CONCLUSIONS

This paper reported a study about the influence of hydrologic information in long term hydrothermal scheduling. The control policy considered combines a deterministic optimization model with an inflow forecasting model, in an open-loop feedback control framework. At each stage in this control policy, the future inflows were foreseen and a deterministic optimization model obtains an operational decision for the next stage.

Three different inflow forecasting models were compared: the long term average inflow, a lag-one periodic autoregressive model and a neuro-fuzzy network model. The effectiveness of each forecasting model was evaluated by the mean and standard deviation values of hydro generation and operational cost. A case study was performed using the hydro plants of the Grand river in Southestern Brazil. The results show that a better inflow foercasting model can reduce operational cost about 5

7. ACKNOWLEDGMENTS

The authors would like to acknowledge the support provided by the São Paulo State Foundation

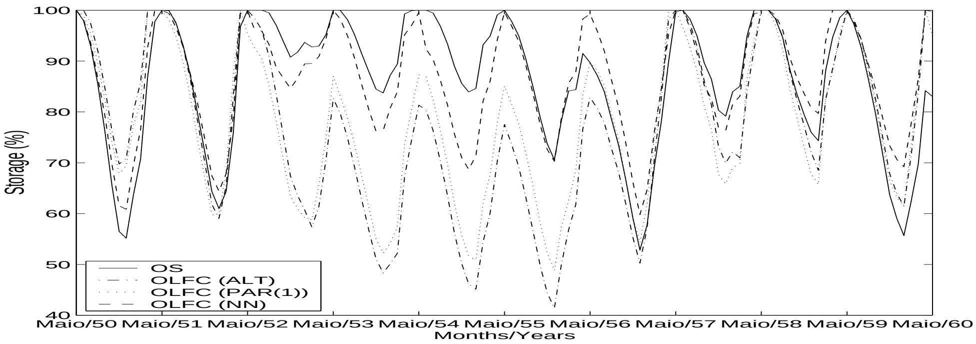

Fig. 2. Hydroelectric test system.

for the Support of Research (FAPESP), and the Brazilian National Research Council (CNPq).

REFERENCES

Ballini, R., M. Andrade and S. Soares (2001). Multi-step-ahead streamflow forecasting by a neurofuzzy network. In: *Joint 9th international Fuzzy Systems Association and 20th North American Fuzzy Information Processing - IFSA/NAFIPS, IEEE.* pp. 992–997.

Bertsekas, D. P. (1995). *Dynamic Programming and Optimal Control.* Athena Scientific. Belmont, Massachusetts.

Bissonnette, V., L. Lanford and G. Côté (1986). A hydro-thermal scheduling model for the Hydro-Quebec production system. *IEEE Transc. on Power Systems* **1**(2), 204–210.

Box, G. E. P., G. M. Jenkins and G. C. Reinsel (1994). *Time Series Analysis - Forecasting and Control.* 3rd ed.. Prentice Hall. New Jersey.

Figueiredo, M., F. Gomide and W. Pedrycz (1995). A fuzzy neural networks: Structure and learning. In: *Fuzzy Logic and Its Applications, Information Sciences and Intelligent Systems* (Z. Bien and K. Min, Eds.). pp. 177–186. Kluwer Academic Publishers. Holanda.

Hanscom, M. L., L. Lafond, L. S. Lasdon and G. Pronovost (1980). Modeling and resolution of the deterministic mid-term energy production problem for Hydro-Quebec system. *Management Science* **26**, 659–688.

Martinez, L. and S. Soares (2002). Comparisons between partial open-loop and closed-loop feedback control policies in long term hydrothermal scheduling. *IEEE Transc. on Power Systems, to appear.*

Oliveira, G. G. and S. Soares (1990). A second-order network flow algorithm for hydrothermal scheduling. *IEEE Transc. on Power Systems* **10**(3), 1635–1641.

Pedrycz, W. and F. Gomide (1998). *An Introduction to Fuzzy Sets: Analysis and Design.* 1rd ed.. MIT Press. Cambridge.

Vecchia, A. V. (1985). Maximum likelihood estimation for periodic autoregressive moving average models. *Technometrics* **27**(4), 375–384.

Yager, R. R. and D. Filev (1994). *Essentials of Fuzzy Modeling and Control.* Wiley Interscience. New York.

Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* **1**, 3–28.