Les Cahiers du GERAD ISSN: 0711–2440

Long-Term Fleet Maintenance of Hydroelectric Generators Using Proportional Hazard Model and Nonlinear Programming

H. Aoudjit, M.-S. Ouali, C. Audet

G–2012–11

March 2012

Long-Term Fleet Maintenance of Hydroelectric Generators Using Proportional Hazard Model and Nonlinear Programming

Hakim Aoudjit
Mohamed-Salah Ouali

Département de mathématiques et de génie industriel
École Polytechnique de Montréal
C.P. 6079, Succ. Centre-ville
Montréal (Québec) Canada, H3C 3A7

hakim.aoudjit@polymtl.ca
msouali@polymtl.ca

Charles Audet

GERAD & Département de mathématiques et de génie industriel
École Polytechnique de Montréal
C.P. 6079, Succ. Centre-ville
Montréal (Québec) Canada, H3C 3A7

charles.audet@gerad.ca

March 2012

Les Cahiers du GERAD

G–2012–11

Abstract

This paper presents a framework to determine optimal maintenance planning of a fleet of complex and independent systems. They are made up of several major components which operate in different environment conditions, and are built with different technologies. This framework uses proportional hazard model (PHM) to characterise the failure rates of components and the effects of the environment conditions and the load levels. A nonlinear programme is developed to minimise the fleet maintenance cost under age replacement policy of its components and a set of organisational and technical constraints. Lindo API and NOMAD are used to solve the nonlinear model. The framework is applied to set a preliminary plan to overhaul a fleet of 90 hydroelectric generators in 6 power plants over 50 years. Sensitivity and performance indexes are built to interpret optimisation results in two settings: normal and 50% increase in load.

Key Words: Maintenance planning, Stochastic Models, Optimization, Nonlinear programming.

1 Introduction

During the end of the last century, several hydroelectric infrastructure investments have been made to provide clean electric power to millions of users across the word, specifically in North America and Asia. These facilities have begun to show serious signs of aging and their maintenance costs to increase significantly over time due to a growing number of hydroelectric generators arriving at the end of their useful life. Maintenance managers apprehend then to face a number of overhauls exceeding what can be handled each year. In addition, increased needs in clean power energy are expected to speed up wear that can lead to early overhaul of many units at the same time.

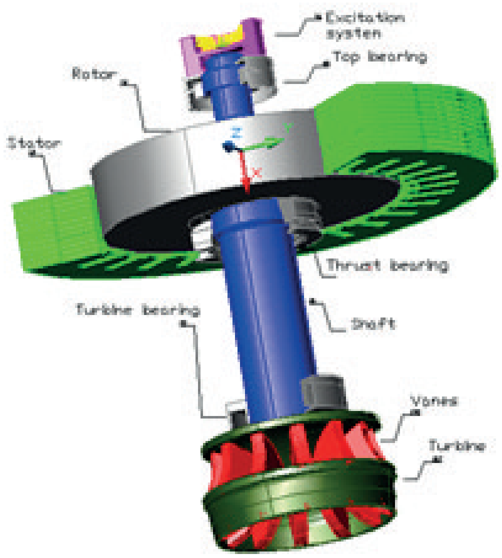

Figure 1: Schematic view of a hydroelectric generator unit

These hydroelectric generator units (HGUs) are complex systems and their main components are functioning in different environmental conditions and are of different technologies (Figure 1). Wear and tear are taking place but for different reasons. While the alternator is subject to heat and electric stress, the turbine is subject to cavitations and corrosion. But still: it takes on average many decades for a newly installed HGU to undergo an overhaul. For such highly reliable machinery, planning replacements many years in advance is not an easy task due to the lack of complete failure observations. Also, standard maintenance policies are difficult to implement regarding the size of the fleet. For example, suppose that alternators are maintained according to the well known age replacement policy where an alternator is immediately replaced at failure or at a fixed age (Barlow et al., 1996). As one power plant could contain as many as 38 HGUs installed roughly the same year, a shortage in maintenance resources can rapidly arise and one cannot apply this standard age replacement policy. The difficulty is worsened when considering that a fleet could contain many hundreds of units and that maintenance managers have to deal with an array of constraints when planning overhauls operations say in a ten years window. In this context, a mathematical programming framework appears to be a suitable tool to tackle this problem.

In this paper we consider the replacement scheduling problem of a fleet of HGUs where each unit's major component is replaced, for sake of simplicity, according to an age replacement policy. Replacement activities involve cost of new component and labour. Due to the logistics involved to overhaul an entire unit, all the major components are replaced/overhauled to "as good as new state" at the same time. So the age threshold for a unit is applied to all major components.

After a literature review of recent work on maintenance scheduling of hydroelectric generators, we begin by decomposing and sorting an HGU's components to highlight the most influential parts in Section 3. Replacement and failure data gathered from maintenance history and utilization profiles computed from production history are then used to derive a failure rate model for the main components. Using the age replacement policy for each component, an average cost function is then derived for the whole unit. These replacement policies at a unit level, constitute the basis for the calculation of an average, overall cost function for the overhauls scheduling problem for the fleet. This problem is considered as nonlinear.

In Section 4 the objective function minimizing the fleet average total maintenance cost per unit of time over an infinite time span is established and the technical and economic constraints are formulated. Important constraints such as the number of specialised crews, number of cranes in a plant, maximum capacity outage for maintenance and annual budget are considered.

Section 5 solves the above nonlinear program using two optimisation libraries: NOMAD, (Le Digabel, 2011) an implementation of the mesh adaptive direct search algorithm, for derirative-free blackbox optimization (Audet and Dennis, 2006, 2009) and Lindo API, (LINDO, 2009) which employs successive linear programming and generalized gradient reduction methods. Several tests are made using NOMAD and/or Lindo API methods to seek the best plan when normal loads are expected and when a 50% increase in load occurs. A case study consisting of 90 HGUs with 3 major components in each unit under an age replacement policy for each component over a 50 years planning horizon is presented. Conclusions and further research alternatives are presented in Section 6.

2 Literature review

Researchers have been studying this kind of problems since the 1960s. They saw similarities with constrained optimization problems and resolution techniques such as linear and dynamic programming have been applied for small thermal production networks. Objectives of existing approaches were levelling reserves (Christiaanse and Palmer, 1972; Garver, 1966; Chen and Toyoda, 1990) or minimizing loss of load probability LOLP during a given year.

The mixed integer linear programming generalizes linear programming by imposing some or all variables are restricted to be integers. The linear nature of the objective function and constraints, allows the development of methods that exploit this structure. However, the fact that some variables are integers, leads to a combinatorial complexity which greatly reduces the size of the problems that can be resolved. Approaches for this class of problems include the implicit enumeration, cutting methods and relaxations techniques (Nemhauser and Wolsey, 1988).

Dopazo and Merrill (1975) have proposed a 0-1 linear integer program for the generator maintenance planning, guaranteeing an optimal solution if it existed. This formulation was dealing with the problems of small size.

In the same period, Zurn and Quintana (1975) showed that dynamic programming with successive approximations, addresses better the size inherent difficulty.

The network interconnection constraints are considered (Marwali and Shahidehpour, 1998) to take account of limitations due to transmission lines. An integrated maintenance plan is formulated and solved by Benders decomposition. The approach is deterministic in (Marwali and Shahidehpour, 1998) and stochastic (Marwali and Shahidehpour, 1999). Stochastic programming is a modeling framework of optimization problems that involve uncertain variables, characterized by probability laws. It attempts to find a feasible solution for almost all of the variables considered while maximizing the expectancy of a certain decision function of random variables. More generally, such models are formulated, solved analytically or numerically, and analyzed to provide useful information for decision making.

The problem of maintenance planning is complex as it deals with not only optimizing an objective function but has also to cope with several types of constraints. A non comprehensive overview is outlined in (Mukerji et al., 1991; Gallestey et al., 2002; Chen et al., 2011).

3 Production profiles and failure analysis

3.1 Production data

Production data of several HGUs from 3 plants located over different river basins has been analysed. These recent HGUs are loaded with electronic devices and sensors which allow huge amounts of data to be gathered for different technical and operational parameters. Production data is provided in the form of the hourly power output over 6 years period. We noted that a typical cycle begins by a unit being placed at a fairly low power output for about an hour (as a warm-up stage), and then the unit is brought to the power output required for the desired period. At the end of a cycle, the decrease in power is also similar (transient stages obviously). Over one year period, production profile fluctuates here between 100 and 120 MW. Cumulative production at each power level is also plotted and shows that there is no production 30% of the time (Figure 2).

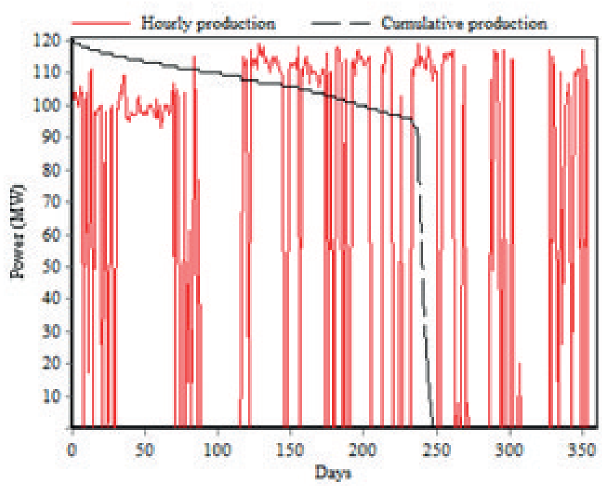

Figure 2: Typical production profile in plant C3

All HGUs do not seem to share the same amount of load. Plots of average load profiles show that plant C3 is more in demand than the other two. Also, simple linear regressions allow us to model with good accuracy a "standard" load to uses as a predictor for failure rate behaviour.

3.2 Failure data

Forced outages data for a sample of 90 HGUs was analyzed and compared to similar historical data described in (Anon, 2003). Although complete failures are observed, they exist only for the oldest plants. Maintenance actions are also reported for 119 components. It appears that only 23 of them participate in 80% of the total unavailability. The largest cumulative maintenance hours (Figure 3) relate to the generator, the turbine, and excitation system. We consider then an HGU as a three components system with an alternator, a turbine and an excitation system denoted respectively by ALT, TURB, and EXS.

Kaplan Meier estimates are used to plot failure rates with censored data for each of the three components. Plots showed clear increasing failure rate for all of them (Figure 4a, b). There is also a variation across plants (Figure 5). Considering this, the proportional hazard model template (PHM) of (Cox and Oakes, 1984) appears to be a good candidate for modeling failure rates.

3.3 Failure rate model

The PHM binds failure rate of a component i of HGU j in an environment that is described by $\boldsymbol{Z}_{ij}$ to a reference state $\boldsymbol{Z}_{ij}^0$ by:

$$\lambda_{ij}\left(t, \boldsymbol{Z}_{ij}\right) = \psi\left(\boldsymbol{Z}_{ij}\right) \lambda_{ij}^0\left(t\right), \tag{1}$$

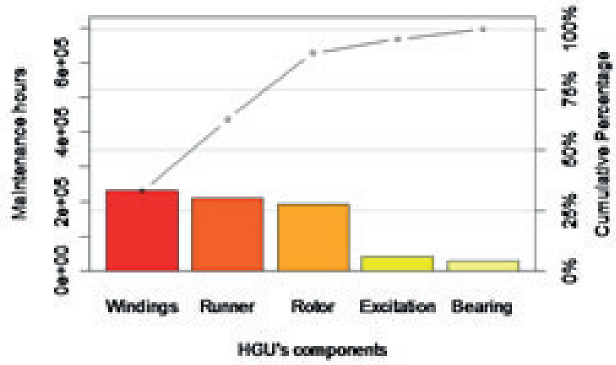

Figure 3: Pareto chart of maintenance hours per component

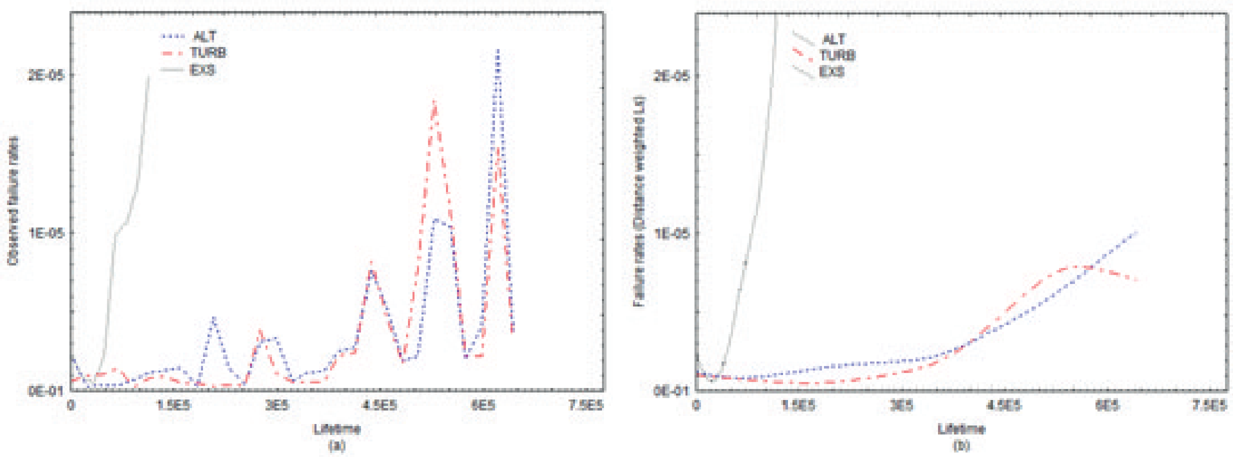

Figure 4: a) Observed, b) distance weighted LS failure rates for components

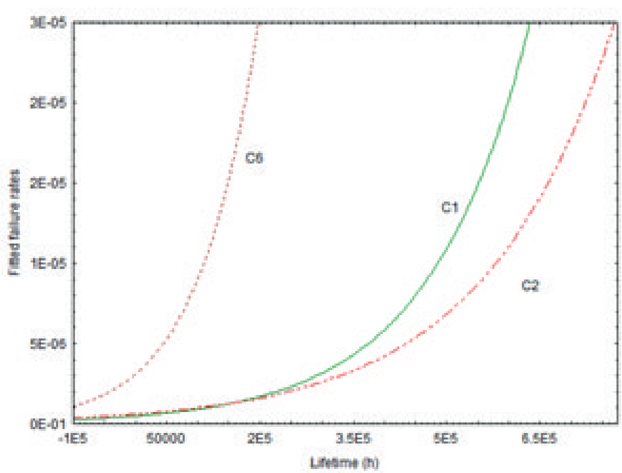

Figure 5: Fitted failure rates for ALTs across plants

where $\psi\left(\boldsymbol{Z}_{ij}^{0}\right) = 1$. This is a widely used approach for modeling a single component system in engineering. The most usual case is where the failure rate is the product of a base failure rate function $\lambda_{ij}^{0}(t)$ and an exponential function designed to express a "relative risk" generated by elements of the vector of covariates

$\boldsymbol{Z}_{ij}$, weighted by coefficients β_{ij}^k so that $\psi\left(\boldsymbol{Z}_{ij}\right) = exp\left(\sum_{k=1}^{r} \beta_{ij}^k z_{ij}^k\right)$. Thus, equation (1) becomes:

$$\lambda_{ij}\left(t, Z_{ij}\right) = e^{\sum_{k=1}^{r} \beta_{ij}^k z_{ij}^k} \lambda_{ij}^0\left(t\right). \tag{2}$$

Using (1), the reliability function takes the following particular form:

$$\frac{f_{ij}\left(t, Z_{ij}\right)}{\overline{F}_{ij}\left(t, Z_{ij}\right)} = \Psi\left(Z_{ij}\right) \frac{f_{ij}^0\left(t\right)}{\overline{F}_{ij}^0\left(t\right)} \Longrightarrow \frac{d\overline{F}_{ij}\left(t, Z_{ij}\right)}{\overline{F}_{ij}\left(t, Z_{ij}\right) dt} = \Psi\left(Z_{ij}\right) \frac{d\overline{F}_{ij}^0\left(t\right)}{\overline{F}_{ij}^0\left(t\right) dt}$$
$$\Longrightarrow ln\left(\overline{F}_{ij}\left(t, Z_{ij}\right)\right) = \psi\left(Z_{ij}\right) ln\left(\overline{F}_{ij}^0\left(t\right)\right)$$

Finally:

$$\overline{F}_{ij}\left(t, Z_{ij}\right) = \overline{F}_{ij}^0(t)^{\psi\left(Z_{ij}\right)}. \tag{3}$$

When $\psi\left(\boldsymbol{Z}_{ij}\right) > 1$, component's reliability will diminish over time compared to the reference state. The model speeds up time as the case where severe usage accelerates ageing.

The main feature of this model is the proportionality between two similar components running in two states $\boldsymbol{Z}_{ij}^1, \boldsymbol{Z}_{ij}^2$. Indeed, the rate:

$$\frac{\lambda_{ij}\left(t, Z_{ij}^1\right)}{\lambda_{ij}\left(t, Z_{ij}^2\right)} = \frac{\Psi\left(Z_{ij}^1\right) \lambda_{ij}^0\left(t\right)}{\Psi\left(Z_{ij}^2\right) \lambda_{ij}^0\left(t\right)} = \frac{\Psi\left(Z_{ij}^1\right)}{\Psi\left(Z_{ij}^2\right)}, \tag{4}$$

is constant. Statistical analysis of failure data, coupled with production data, showed that failures depend heavily on load z_1 and the number starts and stops z_2. For the sake of simplicity, we only consider the load $z = z_1$ as a predictor of failure rates and that the base risk is constant so that:

$$\lambda_{ij}^0\left(t\right) = \vartheta, \tag{5}$$

and that components failure rates take the form:

$$\lambda_{ij}\left\{t, Z_{ij}\right\} = \vartheta e^{\beta_{ij}^1 z} \cdot e^{\beta_{ij}^2 z t}. \tag{6}$$

Parameters are then estimated by comparing the failure rates obtained by a least squares estimation of the observed failure rates.

3.4 Failure rate and production scenarios

We also use the features of model (6) as a tool to forecast wear increase regarding the load: the same component operating today in $\boldsymbol{Z}_{ij}^1$ environment will have a new forecasted failure rate $\boldsymbol{Z}_{ij}^2$ in a new environment of the form:

$$\lambda_{ij}\left(t, Z_{ij}^2\right) = \frac{\Psi\left(Z_{ij}^2\right)}{\Psi\left(Z_{ij}^1\right)} \lambda_{ij}\left(t, Z_{ij}^1\right). \tag{7}$$

Given a future production profile, we have here a simple and direct model to estimate the corresponding forecasted failure rate. According to (7), if the rate $\psi\left(\boldsymbol{Z}_{ij}^2\right) / \psi\left(\boldsymbol{Z}_{ij}^1\right)$ is positive, failure rates are "accelerated" with a certain coefficient in power plant j. As the variability in the failure rate of model (6) is only explained by load, we use failure rate settings of Figure 6 to generate a production scenario.

In this one, accumulated load is increased by 50% ($z_2 = 1.5 z_1$). For example, alternator from plant C2 has an estimated failure rate of the form:

$$\lambda_{21}\left\{t, Z_{21}^1\right\} = 6.1723\ 10^{-7} e^{4.811\ 10^{-6} t},$$

and thus, the failure rate with scenario 2 writes:

$$\lambda_{21}\left\{t, Z_{21}^2\right\} = e^{1.5} \cdot \lambda_{21}\left\{t, Z_{21}^1\right\} = 6.1723\ 10^{-7} e^{7.2165\ 10^{-6} t}.$$

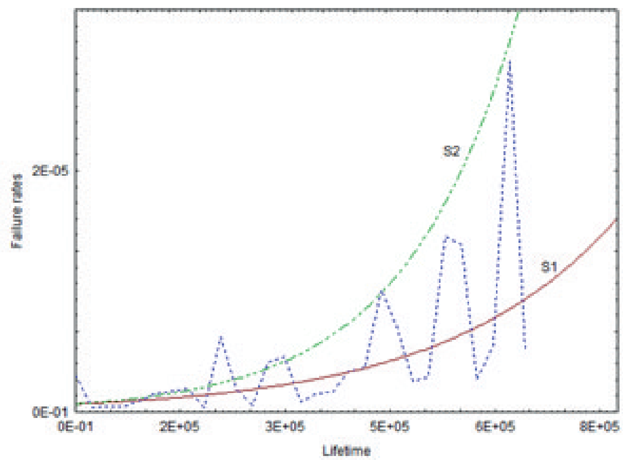

Figure 6: Alternators failure rates with 50% increase in load

4 Fleet maintenance planning

We consider that a fleet of n HGUs are observed over a long period of time and only complete failures are recorded. These failures require a replacement, i.e. the replacement of all major components. All other failures and maintenance without replacements are neglected. Preventively, all major components are replaced according to a periodic age replacement policy (ARP). We seek to find the optimal set of replacement periods of each unit as to minimize the total average maintenance cost per unit of time of the fleet.

We define the fleet periodic replacement policy by a vector τ of n replacement periodicities taking place as the following: a unit in the fleet is replaced preventively at each $k \times \tau_j$, $k = 1, 2, 3 \ldots$ units of time or when a major failure of one of its components occurs. Major failures for a component are described by known probability functions with c.d.f $F_{ij}(\cdot)$. A cost c_{ij}^p is incurred for each component preventive replacement and a cost $c_{ij}^f > c_{ij}^p$ is incurred at its failure.

Assuming that a unit will be indefinitely and preventively replaced, at each τ_j units of time or at failure, and using a renewal reward argument (Ross, 2000), the average total maintenance cost per unit of time $J_i(\tau_j)$ for component i of unit j on an infinite time span is the ratio of expected cost in a replacement cycle by the expected length of a cycle. The average cost in one cycle is given by:

$$C_i(\tau_j) = c_{ij}^f F_{ij}(\tau_j) + c_{ij}^p \overline{F}_{ij}(\tau_j). \tag{8}$$

The average length of a replacement cycle is given by:

$$l_i(\tau_j) = \int_0^{\tau_j} u dF_{ij}(u) + \tau_j \int_{\tau_j}^{\infty} dF_{ij}(u) = \int_0^{\tau_j} \overline{F}_{ij}(u)\, du. \tag{9}$$

One HGU's average maintenance cost per unit of time on an infinite time horizon is then:

$$J_i(\tau_j) = \frac{C_i(\tau_j)}{l_i(\tau_j)}. \tag{10}$$

In a non constrained problem, the minimum replacement cost for a major component is reached for an optimal replacement duration τ_{ij}^{∞} solution of the equation (Barlow et al., 1996):

$$\frac{dJ_i(\tau_j)}{d\tau_j} = 0 \Leftrightarrow \lambda_{ij}(\tau_j) \int_0^{\tau_j} \overline{F}_{ij}(u)\, du - F_{ij}(\tau_j) = \frac{c_{ij}^p}{c_{ij}^f + c_{ij}^p}, \tag{11}$$

where $\lambda(\cdot)$ is the failure rate (Figure 7).

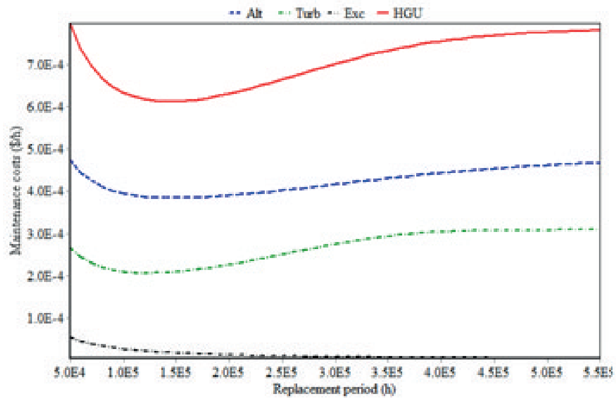

Figure 7: Optimum replacement periods for major components and HGU

Now if we consider that each HGU has I components then the total average maintenance cost per unit time on an infinite time span for unit j is given by:

$$J(\tau_j) = \sum_{i=1}^{I} J_i(\tau_j) = \sum_{i=1}^{I} \frac{c_{ij}^{f} F(\tau_j) + c_{ij}^{p} \overline{F}(\tau_j)}{\int_0^{\tau_j} \overline{F}_{ij}(u)\, du}. \tag{12}$$

HGU optimal preventive replacement period τ_j^*, provided that it exists in $[0, \infty[$, is the solution of

$$\min_{\tau_j} J(\tau_j) \ . \tag{13}$$

Figure 7 shows simulated average total costs per unit of time curves using observed failures rates for each major component and for HGU.

4.1 Maintenance planning objective

The replacement scheduler must decide when and how many units to withdraw from production in a given period of time. Generally, decisions of withdrawals are provided in yearly plans and must contribute to optimize a function on a predefined time horizon. Also, the scheduler must cope with a fleet that can contain hundred of units and an array of constraints.

We consider the fleet of n units and a planning horizon divided into t intervals. At first glance, the size of the problem is $n \times t$ variables. Each HGU has I components. Given the non linearity of the total average maintenance cost per unit time (12), non linear programming (NLP) framework is used here to model the scheduling problem.

The general form of an NLP is:

$$\begin{cases} \min_{\text{s.t}} J(\tau) \\ g_m(\tau) \leq 0, m = 1, 2, \ldots \\ h_l(\tau) = 0, l = 1, 2, \ldots \\ \tau \in \Theta \end{cases} \tag{14}$$

where $J, g_1, g_2, \ldots, h_1, h_2, \ldots$ are functions defined on $\mathbb{R}_n$, Θ a subset of $\mathbb{R}_n$ and τ a vector of n elements. In this problem, the objective is to minimize $J(\tau)$ while fulfilling inequality constraints expressed by g_m and equality constraints h_l. A vector $\tau \in \Theta$ satisfying all constraints is said to be a feasible solution and the set all feasible solutions is the feasible region $\overline{\Theta}$ with $\overline{\Theta} \subset \Theta$. The optimal solution of the problem τ^* is the vector such as$J(\tau^*) \leq J(\tau), \forall \tau \in \overline{\Theta}$.

For the formulation of the NLP planning problem, we consider the previous fleet replacement strategy where ARP is applied to all major components in a unit. If we consider other maintenance policies, the formulation of the planning problem is similar. Recall that under this policy, component i of unit j is replaced when it ages to $\tau_{ij} = \tau_j$, or when a major failure occurs.

Remember that, HGU's total average maintenance cost per unit of time on an infinite time horizon is given by (12). Under (14) framework, elements of vector τ (its dimension equals the number of units to schedule) are the optimal replacements periods for each unit in the fleet under a set of constraints. The objective function is the fleet total average maintenance cost per unit of time computed as the sum of individual unit total average maintenance cost per unit of time:

$$J(\tau) = \sum_{j=1}^{n} J(\tau_j) = \sum_{j=1}^{n} \sum_{i=1}^{I} J_i(\tau_j). \tag{15}$$

The problem (14) is then written as:

$$\begin{cases} \min_{\text{s.t}} \sum_{j=1}^{n} \sum_{i=1}^{I} \frac{c_{ij}^{f} F(\tau_j) + c_{ij}^{p} \overline{F}(\tau_j)}{\int_0^{\tau_j} \overline{F}_{ij}(u) du} \\ g_m(\tau) \leq 0, \ m = 1, 2, \dots \\ h_l(\tau) = 0, \ l = 1, 2, \dots \\ \tau \in \Theta \end{cases} \tag{16}$$

This NLP problem has a special structure because the objective function is called "separable". This facilitates the calculation of the objective function, especially considering its non linear parts. In this case, each portion of the objective function depends only on one variable. If the computation of the objective function is time consuming when iterating through the feasible domain, only the portion of the objective related to the values of the variables that have changed is computed again.

4.2 Maintenance planning constraints

Constraints in overhauls planning appear at various levels of fleet operations. We use the indicator variable x_{jt} to specify status of unit j at time interval t as follow:

$$x_{jt} = \begin{cases} 0 \text{ unit } j \text{ is in overhaul state at period } t \\ 1 \text{ otherwise.} \end{cases} \tag{17}$$

4.2.1 Overhaul start

Specifies for each unit, the period s_j to begin execution of maintenance operations:

$$x_{jt} \in \{0, 1\}, s_j \leq t \leq l_j + d_j. \tag{18}$$

4.2.2 Overhaul duration and continuity

Specifies that the unit is continuously in a shutdown state for the duration of maintenance:

$$x_{jt} = 0, s_j \leq t \leq s_j + d_j. \tag{19}$$

4.2.3 Amortization period

For each category of components, a depreciation period ρ_{ij} is defined. Significant penalties are incurred if the withdrawal is made before this period. We limit the withdrawals beyond this period.

$$x_{jt} = 1, s_j \leq \rho_{ij}, \forall i. \tag{20}$$

4.2.4 Maintenance crews and resources

Overhauling units requires the availability of a multidisciplinary team and resources. In addition this constraint requires that a team is assigned to one unit at the same time. Therefore, let's assume that a unit j requires quantity r_{jk} of resource k for maintenance. Availability of resource k at time interval t is β_{kt}:

$$\sum_{j} (1 - x_{jt}) r_{jk} \leq \beta_{kt}, \forall k. \tag{21}$$

4.2.5 Bulk units

A critical number of units must be available at all times to restart network if a full network shutdown occurs. Withdrawal of a set Φ of units cannot be planned at the same time.

$$\sum_{j \in \Phi} (1 - x_{jt}) \leq 1, \forall t. \tag{22}$$

4.2.6 Maximum capacity outage

No more than a fixed amount of power $\overline{w}_y$ can be shutdown per year y. If w_{jt} is power output of unit j at time interval t then:

$$\sum_{j} (1 - x_{jt}) w_{jt} \leq \overline{w}_y, \forall t \in y. \tag{23}$$

4.2.7 Smoothing maintenance disbursements

This is a constraint stipulating that annual disbursements are of the same order of magnitude. Generally, companies allocate a fixed amount b_y each year to cope with maintenance operations. This corresponds to a disbursement of $\overline{b}_y$ amount per unit of time. It is then necessary to adjust the objective by subtracting a constant:

$$\min J(\tau) \ - \overline{b}_y. \tag{24}$$

4.3 Optimization algorithms

Evaluation of the objective function (16) has a strong nonlinear part consisting of integrals which are numerically computed. It is a numerical code that also evaluates certain constraints. This is an ideal context to the so-called direct search algorithms which uses only functions evaluations to find the best solution. We use the NOMAD library (Le Digabel, 2011) that implements various search algorithms such as MADS (Mesh Adaptive Direct Search algorithm) (Audet and Dennis, 2006, 2009) and whose source code is public. The library also handles biobjective optimisation (Audet et al., 2008). The MADS algorithm is supported by a rigorous hierarchical convergence analysis based on the degree of smoothness and non-smoothness of the objective and constraint functions.

We also use the commercial library LINDO API v5.0 (LINDO, 2009) which offers a multitude of solvers. Moreover, it provides an interface to various programming environments. The standard nonlinear solver of the LINDO API employs successive linear programming (SLP) and generalized gradients reduction (GRG) method. Nonlinear Solver returns a local optimum solution for the problem under consideration. If the local optimal cannot be reached, a feasible solution is reported if it has been found. If a feasible solution cannot be found, or the problem is not bounded or numeric problems have been encountered, a message is returned to the user. The LINDO API can make automatically linear a number of nonlinear constraints by adding other constraints and integer variables, so that the transformed linear model is mathematically equivalent to the nonlinear original one.

For constraints processing, a $n \times t$ matrix initially filled with 1 is built. At each time interval t in the planning horizon where a unit is under overhaul, the corresponding cell is filled with 0. The first time interval where overhaul will take place is obtained using the date of first commissioning and periodicity τ_j. Then subsequent overhaul intervals are filled using overhaul duration. Checking for constraints is then straight forward. Constraints satisfaction is done column-wise (-1 is returned whenever a constraint is met).

5 Application

The planning framework presented here is applied to set a preliminary overhaul plan for a fleet of 90 hydroelectric generators in 6 power plants over the next 50 years time period. In this particular network including more than 300 units, maintenance manager need to prepare for the arrival of a great number of units to the end of their useful life. They also need to know what to expect if there is a shift in load in particular. For this, sensitivity and performance indexes are built to interpret optimisation results under two settings: a first one where load is "standard" as it follows the pattern of past few years and a second one where a 50% increase in load is expected.

In this section a series of 3 problems is solved to plan withdrawals using different contexts. The maintenance policy for all components is the ARP policy. Failure rates used are estimated from failure data history. Because failure rates are increasing, probability density and reliability functions are computed using these failure rates through a Gompertz-Makeham law (Meeker and Escobar, 1998). Each generated plan has also the following metrics (Table 1): 2 average reliability indices are defined on the 50 years planning horizon: an average unit reliability index (AURI) at the time of withdrawal, and the average fleet reliability index (AFRI) defined as the average of reliabilities of all remaining functioning units in the fleet at the time of the successive withdrawals epochs.

To our knowledge no off-the-shelf computer program has the flexibility to achieve the intended planning goals according to modeling involved. We built a computing environment using an SQL database server and a series of dedicated libraries designed to do the job. The machine used is an Intel® Quad Core™ Q93000 workstation at 2.50 GHz, 8 GB of RAM running Microsoft® Windows Vista™ Enterprise edition.

5.1 Assumptions

We need the following additional assumptions to go ahead with optimisation:

(1) Withdrawal of a unit means the replacement of all its components;

(2) Withdrawals history data shows the existence of different installation dates for components for certain units. The date of commissioning of a unit is taken as the oldest component installation date;

(3) As components have different amortization periods, the largest period is met here (50 years = 50 $*$ 8760 hours = 438 000 hours);

(4) Lack of precise maintenance costs records led us to use approximate replacement costs for components: ALT: 15 M$, TURB: 10 M$, EXS: 0.25 M$. We made also the assumption that replacement cost is an average between the preventive and corrective replacement cost. We suppose that failure replacement cost is twice the cost of preventive one for ALTs and 1.5 times for TURBs and EXSs. Therefore, preventive and failure replacement costs for ALTs, TURBs, and EXSs are respectively (10, 20), (8, 12), (0.2, 0.3) M$;

(5) Same maintenance policy is applied to all components of a unit.

In addition, only these constraints are taken into account:

(1) Bound constraints: duration and continuity of each withdrawal in weeks equal to 38. Depreciation periods in years : [ALT, TURB, EXS] $\geq$ [50, 50, 30];

(2) Nonlinear constraints: withdrawn capacity $\leq$ 700 MW per year; Bulk units: withdrawal of maximum of 2 units among (C2, C4, C5) in the same interval; Human resources: crews $\leq$ 3; Number of simultaneous withdrawals in each plant limited by lifting capabilities $\leq$ 2; Annual maintenance disbursements $\leq$ 250 M$.

5.2 Optimisation starting point

A starting plan to search for an optimal solution (Figure 8a) is based on replacement periods according to depreciation periods. It is the longest period (50 years) in this case. As expected, this plan is not feasible.

At least two constraints are violated: number of maintenance crews available at each epoch and lifting capabilities in each plant.

5.3 Problem I

This first resolution is aimed at minimizing the total average maintenance cost per unit of time for a fleet of 90 HGUs with a standard production (scenario 1). The resolution is first attempted with the NOMAD Solver. After two attempts with successive increases in the number of function evaluations (10^4, 3.53 min) then (510^5, 9.23 min), no feasible solution is found. The optimization is then restarted with Lindo API v5.0 Solver. The resolution ends with status 5 describing a feasible yet not optimal solution (status 1 for optimal and 8 for locally optimal).

Although starting solution is not feasible, optimization has produced a feasible solution. The generated plan shows that unit's average reliability during its withdrawals is 3.20% and that the rest of the units still in operation are at 35.71%. These indices indicate that units are removed at the limit of the complete failure. Their production potential is used at maximum.

This feasible solution is then taken as starting point for the NOMAD algorithm. An additional gain is obtained on the objective function (Table 1, Figure 8b). It is improved by 143 \$/h for a total average cost per unit time of 7566 \$/h. This gain is however obtained at the cost of 50000 function evaluations in 33 minutes (Figure 9).

Table 1: Features of starting and withdrawals plan for problem I

		Target	Start	PbI Lindo	PbI NOMAD
Objective	$\boldsymbol{J}(\tau)$(k\$/h)	Min	7.712	7.709	7.566
Constraints	Overhaul duration (weeks)	$= 38$	-1	-1	-1
	Amortization period (years)	$\geq [50, 50, 30]$	-1	-1	-1
	Maintenance crews	≤ 3	19	-1	-1
	Number of cranes in plant	≤ 2	1	-1	-1
	Bulk units		1	-1	-1
	Max. capacity outage (MW)	≤ 700	1001	-1	-1
	Annual disbursements (M\$)	≤ 250	67.56	67.53	66.28
Indexes	AURI (%) 50 years	–	26.60	3.20	9.82
	AFRI (%) 50 years	–	66.79	35.71	57.52
	Average age at withdrawal (yrs)	–	50	81	66
	Resolution time	–	–	3.92	33.00

Figure 9 show that NOMAD made most gains on early 20000 evaluations. Subsequently, it appears to stagnate and the algorithm doesn't make more major improvements. If resolution time is a factor to be considered for a greater number of units, it would be necessary to find a compromise between function evaluations and objective improvements.

The plan has the following additional features:

(1) A total of 38 additional units appear in the 50-year plan. Withdrawals are better spread along the time horizon;

(2) The units are not used to full potential. The average reliability of units at their withdrawal is greater and reaches 9.82%;

(3) Average reliability of units remaining in operation is also greater at 57.52%.

This plan is the best withdrawal plan on 50 years with scenario I so far.

5.4 Problem II

The issue considered here is to illustrate the effect on withdrawal planning with a 50% load increase (scenario 2). We use the best plan of problem I as the resolution starting point to show the changes that occur

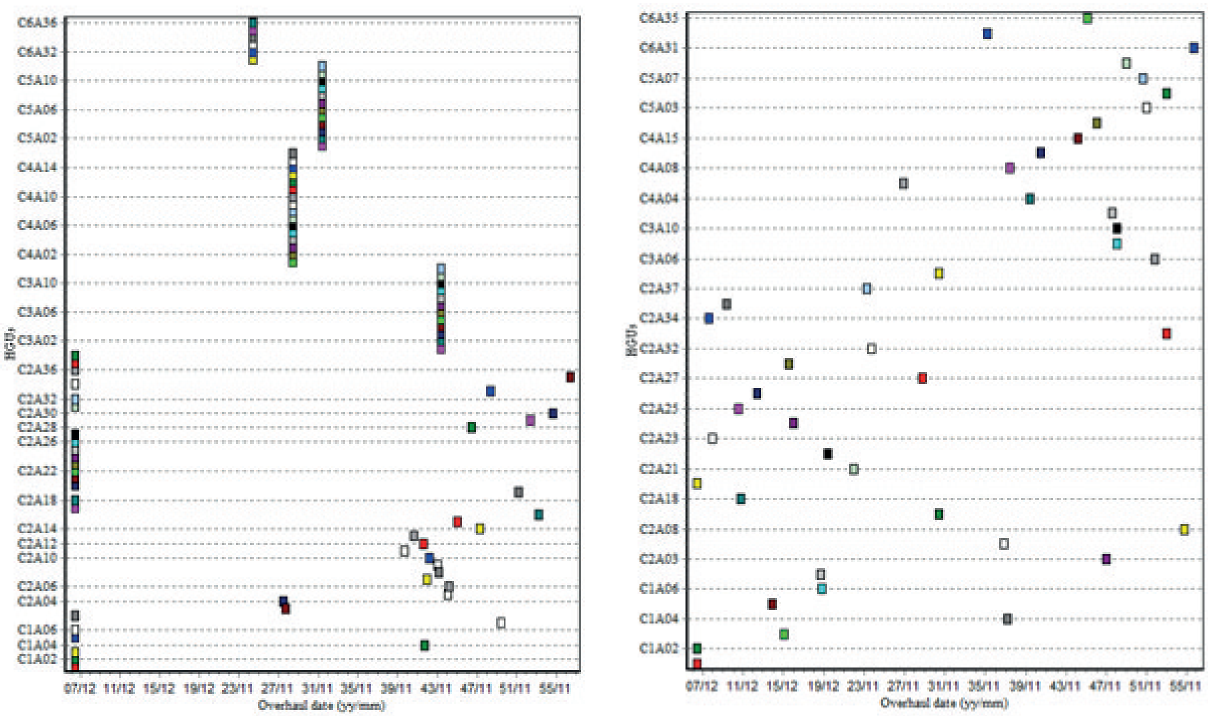

Figure 8: Diagrams of a) starting and b) best withdrawals plan for problem I

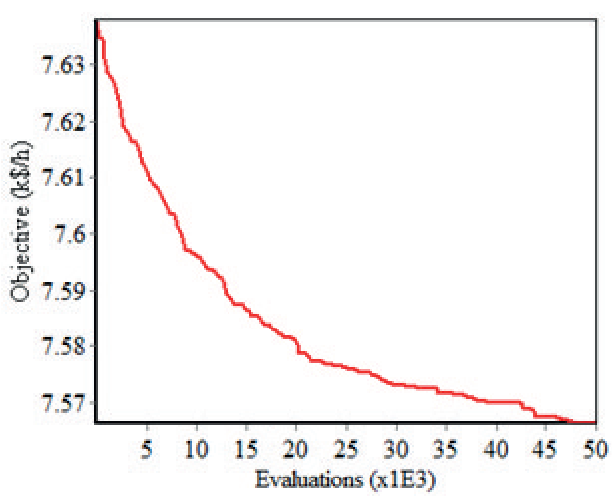

Figure 9: Problem I objective improvement using NOMAD solver

on a plan that has already been implemented for example. Successively, the NOMAD solver and LINDO API v5.0 are launched with no luck for NOMAD after 10000 evaluations. Lindo get a feasible plan in 17 min (Table 2).

This plan has the following features:

Table 2: Features of withdrawals plan for problem II

		Target	PbII Lindo	PbII NOMAD
Objective	$\boldsymbol{J}(\tau)$(k\$/h)	Min	9.799	9.562
Constraints	Overhaul duration (weeks)	$= 38$	-1	-1
	Amortization period (years)	$\geq [50, 50, 30]$	-1	-1
	Maintenance crews	≤ 3	-1	-1
	Number of cranes in plant	≤ 2	-1	-1
	Bulk units		-1	-1
	Max. capacity outage (MW)	≤ 700	-1	-1
	Annual disbursements (M$)	≤ 250	85.84	83.76
Indexes	AURI (%) 50 years	–	0.97	10.58
	AFRI (%) 50 years	–	31.69	52.35
	Average age at withdrawal (yrs)	–	71	55
	Resolution time	–	6.17	28.40

Similarly, the same workaround is used to improve the objective by NOMAD. An additional gain is also found on the objective function (it is enhanced by 237 $/h). The average total cost per unit of time is 9.562 k$/h. The plan has also the following additional features:

(1) All 90 units in the fleet appear in the 50-year plan. Withdrawals are still better distributed on the horizon;

(2) HGUs are not at end of life. AURI is greater at 10.58%;

(3) AFRI is also increased at 52.35%.

By comparing plans generated for problems I and II, Table 3 shows that withdrawals events are premature by an average of ten years when doubling the load. Although the AURI remains stable, AFRI for the remaining operating units is lower by 5.17%. To maintain the same AFRI as for scenario 1, one needs to relax a constraint (by augmenting crews for example).

Table 3: Comparison of features of withdrawals plans for problems I and II

		Scenario 1	Scenario 2	Variation
Objective	$\boldsymbol{J}(\tau)$(k\$/h)	7.566	9.562	1.995
Indexes	AURI (%) 50 years	9.82	10.58	0.76
	AFRI (%) 50 years	57.52	52.35	-5.17
	Average age at withdrawal (yrs)	66	55	-11
	Resolution time NOMAD (min)	33.00	28.40	-4.6

6 Conclusions

In this paper, nonlinear mathematical programming methodology was applied to prepare the preliminary maintenance planning of a fleet of hydro generators units. It provides hydroelectric plants managers with a support tool to forecast their assets withdrawals based on production data and maintenance policies years ahead. The mathematical program was designed to assess an objective function according to the choice of a maintenance policy for each component in a unit, and under various types of constraints including technological, economic and human resources restrictions. Moreover, it was shown how to use proportional hazard model to predict accelerating degradation when load was considered to be the most influential parameter in failure rate variation. The impact of usage on withdrawal plan is then introduced through the objective function. A simulated increase of 50% in load has shown a shift in withdrawal plan. They happen 10 years earlier on average in the application settings.

To describe the "performance" of a plan, a set of reliability based indexes have been introduced to compare resulting plans. One can use them to explore ways to improve the performance of a plan by changing the state of a constraint. One can then seek a better index for reliability of remaining operating units by adding more maintenance resources for example.

Although this methodology has been applied to planning withdrawals of a sample of 90 units on a 50 year time horizon, it is quite clear that the proposed developments and results clearly demonstrate the feasibility of such a program for more hydroelectric power units. The information system developed specifically to resolve this scheduling problem is limited only by the ability of nonlinear solvers. Then, as more data is gathered on remaining units, the size of the fleet will be augmented to encompass all units in future work. Also, more general maintenance policies, and costs structure will be introduced and tested.

References

Audet C. and Dennis J.E., 2006. Jr. Mesh adaptive direct search algorithms for constrained optimization. *SIAM Journal on Optimization*, 17(1), 188–217.

Audet C. and Dennis J.E., 2009. A progressive barrier for derivative-free nonlinear programming. *SIAM Journal on Optimization*, 20(1), 445–472.

Audet, C., Savard, G. and Zghal, W., 2008. Multiobjective optimization through a series of single-objective formulations. *SIAM Journal on Optimization*, 19(1), 188–210.

Anon, 2003. Rapport annuel – L'tat des quipements de production. In: Association canadienne de l'lectricit.

Barlow, R.E., Proschan, F., and Hunter, L.C. (1996). *Mathematical theory of reliability.* Philadelphia: SIAM.

Chen, L.-N., and Toyoda, J., 1990. Maintenance scheduling based on two-level hierarchical structure to equalize incremental risk. *IEEE Transactions on Power Systems*, 5, 1510–1516.

Chen, N., Chen, Y., Li, Z., Zhou, S., and Sievenpiper, C., 2011. Optimal variability sensitive condition-based maintenance with a Cox PH model. *International Journal of Production Research*, 49(7), 2083–2100.

Christiaanse, W.R., and Palmer, A.H., 1972. A technique for the automated scheduling of the maintenance of generating facilities. *IEEE Transactions on Power Apparatus and Systems*, PAS-91, 137–144.

Cox, D.R., and Oakes, D., 1984. *Analysis of Survival Data.* Cambridge: Chapman and Hall Ltd.

Dopazo, J.F., and Merrill, H.M., 1975. Optimal Generator Maintenance Scheduling using Integer Programming. *IEEE Transactions on Power Apparatus and Systems*, 94(5), 1537–1545.

Gallestey, E., Stothert, A., Antoine, M., and Morton, S., 2002. Model predictive control and the optimization of power plant load while considering lifetime consumption. *IEEE Transactions on Power Systems*, 17, 186–191.

Garver, L.L., 1966. Effective load carrying capability of generating units. *IEEE Transactions on Power Apparatus and Systems*, PAS-85(8), 910–919.

Le Digabel, S., 2011. Algorithm 909: NOMAD: Nonlinear Optimization with the MADS algorithm. *ACM Transactions on Mathematical Software*, 37(4), 44:1–44:15.

LINDO, S., 2009. LINDO API. In: LINDO Systems (5 ed.).

Marwali, M.K.C., and Shahidehpour, S.M., 1998. Deterministic approach to generation and transmission maintenance scheduling with network constraints. *Electric Power Systems Research*, 47, 101–113.

Marwali, M.K.C., and Shahidehpour, S.M., 1999. Probabilistic approach to generation maintenance scheduler with network constraints. *International Journal of Electrical Power and Energy System*, 21, 533–545.

Meeker, W.Q., and Escobar, L.A., 1998. *Statistical methods for reliability data.* New York: Wiley.

Mukerji, R., Merrill, H.M., Erickson, B.W., Parker, J.H., and Friedman, R.E., 1991. Power plant maintenance scheduling: Optimizing economics and reliability. *IEEE Transactions on Power Systems*, 6, 476–483.

Nemhauser, G.L., and Wolsey, L.A., 1988. *Integer and Combinatorial Optimization.* New York: Wiley.

Ross, S.M., 2000. *Introduction to Probability Models.* San Diego: Harcourt/Academic Press.

Zurn, H.H., and Quintana, V.H., 1975. Generator maintenance scheduling via successive approximations dynamic programming. *IEEE Transactions on Power Apparatus and Systems*, 94(2), 665–671.