Maintenance Strategy Optimization for Complex Power Systems Susceptible to Maintenance Delays and Operational Dynamics

Hindolo George-Williams, and Edoardo Patelli

Abstract—Maintenance is a necessity for most multi-component systems, but its benefits are often accompanied by considerable costs. However, with the appropriate number of maintenance teams and a sufficiently tuned maintenance strategy, optimal system performance is attainable. Given system complexities and operational uncertainties, identifying the optimal maintenance strategy is a challenge. A robust computational framework, therefore, is proposed to alleviate these difficulties. The framework is particularly suited to systems with uncertainties in the use of spares during maintenance interventions, and where these spares are characterised by delayed availability. It is provided with a series of generally applicable multi-state models that adequately define component behaviour under various maintenance strategies. System operation is reconstructed from these models using an efficient hybrid load-flow and event-driven Monte Carlo simulation. The simulation's novelty stems from its ability to intuitively implement complex strategies involving multiple contrasting maintenance regimes. This framework is used to identify the optimal maintenance strategies for a hydroelectric power plant and the IEEE-24 RTS. In each case, the sensitivity of the optimal solution to cost level variations is investigated via a procedure requiring a single reliability evaluation, thereby reducing the computational costs significantly. The results show the usefulness of the framework as a rational decision-support tool in the maintenance of multi-component, multi-state systems.

Index Terms—Monte Carlo Simulation, Complex System, Uncertainty, Maintenance Optimization, Multi-State System.

NOTATIONS

$\boldsymbol{A} - \boldsymbol{B}$	elements in $\boldsymbol{A}$ but not in $\boldsymbol{B}$.
$\lceil a \rceil$	smallest integer greater than a.
$min\,(\boldsymbol{A})$	the least element of set/vector $\boldsymbol{A}$.
$min\,(\boldsymbol{A}, b)$	the least element of $\boldsymbol{A}$ greater than b.
$Exp\,(a)$	exponential distribution with rate $1/a$.
$U\,(a, b)$	uniform distribution with bounds on a, b.
$LogN\,(a, b)$	log-normal distribution with mean a, std. b.
$Wb\,(a, c)$	Weibull distribution with scale parameter a and shape parameter, c.
$Gu\,(a, b)$	Gumbel distribution with mean a, std. b.
$G\,(a, b)$	gamma distribution with shape parameter a and scale parameter, b.
$u \sim [0, 1]$	uniform random number between 0 and 1.
$[a, b]$	maint. strategy based on regimes a and b.
$numel\,(\boldsymbol{A})$	number of elements in set/vector $\boldsymbol{A}$.

The authors are with the Institute for Risk and Uncertainty Engineering, University of Liverpool, UK. Email: H.George-Williams@liverpool.ac.uk (Hindolo), epatelli@liverpool.ac.uk (Edoardo)

Hindolo George-Williams also works with the Institute of Nuclear Engineering and Science, National Tsing Hua University in Hsinchu, Taiwan.

ABBREVIATIONS

APM	Awaiting Preventive Maintenance state.
CM	Corrective Maintenance state.
EENS	Expected Energy Not Supplied.
$(EENS)_{eff}$	total EENS.
D	Diagnosis state
F	Failed state.
I	Idle state
PF	Partial Failure state
PM	Preventive Maintenance state.
S	Shutdown state.
W	Working state.

NOMENCLATURE

p_i	probability of spares for CM of component i.
q_i	probability of spares for PM of component i.
t_{pm}	preventive maintenance duration.
k_i	proportion of t_{pm} spent before spares request.
Λ_i	minimum threshold load for component i.
ω	number of maintenance groups.
n_{t_j}	total number of teams in group j.
n_{1_j}	number of CM teams in group j.
n_{2_j}	number of PM teams in group j.
$\boldsymbol{n}^*$	a combination of maintenance teams
m_j	total number of components in group j.
M	total number of maintainable components.
M'	total number of system nodes.
T_m	mission time.
$\boldsymbol{T}_i$	transition matrix for component i.
N	Number of Monte Carlo samples.
$\mathbb{N}$	Set of possible maintenance team combinations
$N_i^{\{cm\}}$	number of CM actions on component i.
$N_i^{\{pm\}}$	number of PM actions on component i.
$t_i^{\{cm\}}$	time spent by component i in CM.
$t_i^{\{pm\}}$	time spent by component i in PM.
$s_i^{\{cm\}}$	number of CM spares used for component i.
$s_i^{\{pm\}}$	number of spares used in PM of component i.
$\mu_i^{\{cm\}}$	CM suspension indicator for component i.
$\mu_i^{\{pm\}}$	PM suspension indicator for component i.
t_s	current simulation time.
x	current state.
y_{next}	next transition state.
t_{next}	next transition time.
y_m	next maintenance state.

y'	next failure state of a component in APM.
t'	maximum lifetime of a component in APM.
t_{spent}	time spent in PM before maint. suspension.
t_{spare}	spares delay time.
t_{rem}	remaining lifetime of a component.
$\theta_j^{\{cm\}}$	set of components repaired by group j.
$\theta_j^{\{pm\}}$	set of components assigned to group j for PM.
θ_j	$\left(\theta_j^{\{cm\}} \cup \theta_j^{\{pm\}}\right)$
$\lambda_j^{\{cm\}}$	number of busy CM teams from group j.
$\lambda_j^{\{pm\}}$	number of busy PM teams from group j.
Π	matrix defining the number of maint. teams.
φ	shared/dedicated maintenance indicator.
h_1	set of components in CM queue.
h_2	set of components in PM queue.
h_{1f}	final content of h_1 after normalization.
h_{2f}	final content of h_2 after normalization.

I. INTRODUCTION

OWING to the rapid growth in human population and the proliferation of new electrical energy-driven technologies, the demand for sustainable electricity is on a steady rise. Coupled with a competitive market, the electrical power operator is under increasing pressure to deliver an adequate, safe, affordable, and uninterrupted supply. They, however, are constrained by the impossibility to continuously operate the system without outages, consequent of component failure and maintenance. To minimize the impact of these outages on consumer satisfaction, the maintenance strategy adopted should be robust, meet operator expectation, extend the life of the system, and be carefully executed [1], [2]. From an operator perspective, a robust strategy is one that ensures maximum system throughput and keeps the operating cost to a minimum. In addition to its impact on system performance, maintenance accounts for a significant proportion of the total operating cost of power systems. It, therefore, to a significant extent, defines the revenue generated and the overall investment sustainability. In summary, the principles of modern maintenance engineering do not only require meeting technical and operational goals, but achieving them through the most cost effective means. This constraint dictates, maintenance follow a strategy imposing minimum system output loss and at the least possible cost.

A. Maintenance Strategy Optimization

In the most general sense, maintenance can be optimized against various reliability and performance indices. The indices used depend on the application and the goal of the analyst. For instance, in nuclear and other safety-critical systems, failure probability and recovery likelihood are the most frequently used indices. However, regardless of the application and the indices used, the goal is, finding the optimum balance between costs and benefits, whilst not ignoring any important system constraints [2]. This process involves comparing the monetary equivalent of the benefits, to the costs incurred in their attainment. A limiting factor, therefore, would be the convertibility to monetary gains, of these benefits. Consequently, cost minimization has been the subject of many maintenance

optimization models [1], [3]–[12]. While some of these models consider the system as a single unit (for instance [1], [6], [13]), many are enhanced for multi-component systems. With respect to implementation effort, multi-component models are more demanding, due to the presence of multiple system dynamics and structural complexities. Notwithstanding, various researchers have successfully implemented maintenance optimization models on multi-component systems [3]–[5], [8]–[11]. A comprehensive review and historical overview can be found in [14]–[16].

The cost of maintaining a system constitutes various parameters, varying according to the external dynamics surrounding the system and the intrinsic properties of its building block. Prominent amongst these are the reliability and maintainability of components, cost of spares, labour cost, and the frequency and duration of preventive maintenance actions. An accurate model, therefore, accounts for all of these parameters. With a few exceptions focusing on reliability-centered maintenance [5], [8] or maintenance contract assessment [17], most models are dedicated to determining either the optimal preventive maintenance schedule, inspection, or component replacement intervals. Often, they are hinged on the assumption that there are sufficient maintenance teams to accomplish maintenance functions, [4]–[9], [11], [17], and delays imposed by logistic and administrative constraints are usually ignored [3]–[9], [11], [17]. Instantaneous preventive maintenance or inspection is another assumption frequently invoked [3], [4], [9], [11], [13]. While these assumptions are reasonable for some systems, they may be completely unrealistic for many. A notable instance being, a system with large maintenance durations and operated under limited maintenance team conditions. These large durations, normally due to logistic or human resource constraints, affect system performance negatively. They also render the cost and number of spares used worth considering, a factor many maintenance optimization models have ignored.

When the possibility of maintenance interruptions exists, constraints on the states of components during periods of maintenance suspension become important. A component's maintenance is suspended if it requires spares which availability is delayed or if the maintenance team is reassigned to a more critical component. During suspensions, the component may either be put back into operation (assuming it is only partially failed or under preventive maintenance) or kept out of operation until maintenance is completed. The careful scheduling of these maintenance actions may also mitigate their effect on throughput losses. This is the case, especially for planned preventive maintenance and corrective maintenance of partially failed components. Hence, there is the need for an optimization framework that derives the combination of procedures (maintenance strategy) minimizing system losses, as well as the maintenance cost. Maintenance strategy here refers to a set of procedures specifying,

1) The number of maintenance teams employed and how they are assigned to components.
2) Whether or not preventive and corrective maintenance should be carried out by the same team.
3) Whether preventive maintenance interventions and corrective maintenance of partially failed components

should consider the state of the system or a relevant subsystem.

4) What happens to a component when its maintenance is suspended.

Significant strides have already been made toward maintenance strategy optimization in the presence of some of these, including other dynamic considerations like ageing, imperfect and condition-based maintenance [3], [4], [18], [19]. However, the techniques proposed in these works are suited mainly to binary-state systems. An approach considering all the constraints in question and in a multi-state, multi-component environment is yet to emerge. In this work, a simulation framework that can be used to identify the optimal maintenance strategy for a multi-state system prone to the range of possible operational dynamics listed is proposed. A detailed account of its theoretical and modelling principles is provided, thereby setting the tone for its wide applicability.

B. Advantages of the Proposed Approach

The dependability of the optimal solution obtained from any maintenance strategy optimization scheme is determined by the accuracy of its system performance measures. This, in turn, is influenced by the suitability to the system, of the reliability modelling technique employed. These modelling techniques fall into one of two broad categories; analytical and dynamic reliability models. The former are inapplicable to certain reliability problems, especially those involving complex maintenance strategies and other dynamic considerations. When forced to suit such problems, the resulting models are often oversimplified to an extent that compromises the credibility of the outcome. In fact, most of the limitations of the existing maintenance optimization models discussed in the preceding section are associated with analytical models.

Dynamic reliability models, on the other hand, possess sufficient flexibility to model the dynamic considerations and uncertainties that normally characterize the operation of realistic systems. Stochastic Petri Nets [20], Stochastic Hybrid Systems [21], and Monte Carlo Simulation [3], [22]–[24] are the most popular in this category. Stochastic Petri Nets, however, require the enumeration of the entire state space of the system, which makes them infeasible for complex multi-state systems, even of moderate size. They also suffer a serious setback when the system can undergo non-Markovian transitions, in which case Tuffin et al. [25] recommend simulation. Stochastic Hybrid Systems are an emerging modelling formalism with promising prospects for dynamic reliability modelling. They are built around the Markov reward model of the system, when solicited for problems involving performance optimization or system operating cost minimization [21]. Consequently, like Stochastic Petri Nets, they are intractable for complex multi-state systems, due to their susceptibility to the state explosion conundrum. In addition, they proceed by translating the dynamic reliability problem into a set of differential equations, which closed-form solution, in some cases, may be difficult to obtain analytically. Some researchers [26] have even had to resort to a Monte Carlo Simulation approach to solving these differential equations. Given the structural complexity of most

power systems and their multi-state attributes, Monte Carlo Simulation, therefore, remains the most feasible approach, regardless of its higher computational intensity.

However, most Monte Carlo simulation algorithms [23], [27], [28] require prior knowledge the system's structure function or its path or cut set, which for complex multi-state systems is tedious. In [22], a simple load flow-based simulation approach, applicable to any system configuration was introduced. It allows the simulation of a multi-state system without the need to define its structure function, path or cut sets. Notably, it enables the replication of realistic system operating principles like, shut down and restart of components. These shut down events can be as a result of the unavailability of another component or loading restrictions imposed on the components themselves. When dealing with maintainable systems, it's vital to consider this form of functional interdependency between components, as the failure and preventive maintenance of most components depend on the effective time spent in operation. Most reliability and performance analysis approaches disregard this feature because, it's either impossible or difficult to determine the actual flow through system components. We adapt this modelling approach to systems with limited maintenance teams and prone to maintenance delays and other operational uncertainties. The modified approach is a credible pathway via which system performance indices relevant to the maintenance model are derived, without making unrealistic assumptions.

Appreciating that most power systems exhibit multi-state characteristics, each system component is modelled as a semi-Markov stochastic process. The multi-state model is modified to incorporate additional stochasticity induced by the operational dynamics surrounding the system. Thus, the resultant component model is also a translation of system dynamics, from the system to the component level. This model simplifies the simulation procedure, rendering it more intuitive and generally applicable. Most importantly, the simulation procedure supports the complex scenario where various components follow different maintenance strategies; another limitation of existing models.

The remainder of this work is organised as follows; Section 2 is dedicated to defining key terms, presenting a general overview of the problem under consideration, the proposed cost model, and a description of the solution procedure. In Section 3, a background to the component and system models is presented. The simulation algorithm and details on how components are modelled to account for various system dynamics are also described here. Section 4 presents two case-studies, illustrating the application of the models developed to realistic systems. Finally, in Section 5 a conclusion is drawn on the proposed framework, with insights on its applicability.

II. PROBLEM FORMULATION

Consider a multi-component system of arbitrary structure, composed of either binary-state components, multi-state components, or both. These components can undergo corrective maintenance when in a degraded state and preventive maintenance which interval is determined by the effective time

spent in operation since the last maintenance action (i.e., periods when the component is unavailable do not count). State transition times of components may be constant or follow any probability distribution. On entering a degraded state, a component is added to the maintenance queue and its repair process follows two stages; a diagnosis stage and a stage dedicated to actual repairs. At the end of diagnosis, the maintenance team may proceed to the second stage or initiate a spares request, if spares are required. The probability of the latter happening is p_i, where i, a positive integer arbitrarily assigned, is the index of the component in the system. There's a delay between initiation of spares request and their delivery, which may vary from component to component and may again follow any probability distribution. Like corrective maintenance, preventive maintenance is prone to interruptions at a probability, q_i. This normally is realised after an average time $k_i t_{pm} \mid 0 < k_i < 1$, t_{pm} being the component's expected preventive maintenance duration, and k_i, the proportion of this time to elapse before the need for spares is realised. Whilst the crew awaits the spares, they can be assigned to another job, if there are no other idle maintenance teams.

At the system level, components are arranged into ω maintenance groups, and each group maintained by $n_{t_j} \mid j = 1, 2, ..., \omega$ maintenance teams. Under dedicated maintenance, n_{t_j} is expressed in the form, $(n_{1_j}, n_{2_j}) \mid n_{1_j} + n_{2_j} = n_{t_j}$, where n_{1_j} is the number of teams dedicated to corrective maintenance, and n_{2_j}, the number of teams dedicated to preventive maintenance. It's assumed each of these n_{t_j} teams has the expertise to maintain any of the m_j components in group j. Maintenance is outsourced, and its cost constitutes three parts; a fixed cost per unit time per maintenance team, a fixed cost per maintenance call, and a fixed cost per unit time of actual maintenance service. There are no penalty costs on the system operator for failing to meet demand, but consumers only pay for the quantity of output supplied. The lost revenue accrued, with the total maintenance cost over a period, provides a measure of the performance of the system for that period. It's desired to find the maintenance strategy and the value of $n_{t_j} \forall j \in \{1, 2, ..., \omega\}$, ensuring optimum system performance. The objective of the optimization procedure, is the minimization of system maintenance cost, as well as the cost incurred from unmet demand. A given strategy, therefore, is optimal if it minimizes the total cost.

There are a few attributes of the system described that pose some challenges. From a modelling point of view, the fact that the system could be multi-state and of any architecture, disqualifies most of the existing system reliability evaluation techniques (see Section I-B). Similarly, the limited number of maintenance teams, the uncertainties associated with the need for spares to complete a maintenance action, and the delays in the availability of these spares, present a serious planning and scheduling dilemma. For instance, if the maintenance crew knew every preventive maintenance action would require spares, they would place a spares request in advance. Conversely, they could carry with them a few spares in anticipation, but this would be applicable only to non-bulky components, since there is a limit to how much could be carried. The need, therefore, for an optimal maintenance

strategy cannot be overemphasised.

A. Definition of key terms

1) *Expected Output-not-supplied*: A measure of the expected amount by which the actual system output deviates from its expected level, within a given period, T_m. This quantity, in power systems, is known as the *Expected Energy Not Supplied (EENS)*, and it accounts for the periods the system performance curve is below the load curve. If $Y(t)$ and $Y_d(t)$ respectively denote the instantaneous system output and demand, then, for a demand-driven system (i.e., $Y(t) \leq Y_d(t)$),

$$EENS = \int_0^{T_m} (Y_d(t) - Y(t)) \, dt \qquad (1)$$

For a given system reliability problem, $Y_d(t)$ is normally known, and $Y(t)$ is computed from the system reliability analysis outcome. When obtained via Monte Carlo simulation, $Y(t)$ is defined by a collection of discrete sets of system performance levels, as a function of time. Therefore, the discrete form of (1) should be used to compute the system EENS. Given $Y(t)$ is random, the EENS is computed as the average of the performance deficiencies of all the samples of $Y(t)$. For N Monte Carlo samples of $Y(t)$, let the ith sample contain n_i performance level transitions, $y_{ij} = Y_d(t) - Y(t)$ at the jth transition, and $t = t_{ij} \mid 0 \leq t_{ij} \leq T_m$; the corresponding transition time, then,

$$EENS = \frac{Y_0}{N}$$
$$Y_0 = \sum_{i=1}^{N} \left(y_{in_i} (T_m - t_{in_i}) + Y_1 \right) \qquad (2)$$
$$Y_1 = \sum_{j=2}^{n_i} y_{i(j-1)} \left(t_{ij} - t_{i(j-1)} \right)$$

where y_{in_i} and t_{in_i} are respectively the final performance level and last transition time of sample i. Alternatively, if instead of $Y(t)$ and $Y_d(t)$, only the possible system performance and demand levels with their corresponding occurrence probabilities are known, the EENS is computed through a different approach. Let the system exist in n distinct output levels as defined by vector $\mathbf{C}$, with probability of occurrence within the period, T_m, defined by vector $\mathbf{P}$. The expected performance deviation per unit time, β, and $EENS$ are,

$$\beta = \sum_{j=1}^{\alpha} (j, \mathbf{P}_d) \beta_0^{\{j\}}$$
$$\beta_0^{\{j\}} = \sum_{i=1}^{n} max \left((j, \mathbf{C}_d) - (i, \mathbf{C}), 0 \right) (i, \mathbf{P}) \qquad (3)$$
$$EENS = T_m \beta$$

where α is the number of possible demand levels, $\mathbf{C}_d$, the vector defining these levels, and $\mathbf{P}_d$, the vector specifying their corresponding probabilities of occurrence. For systems like power distribution networks with

multiple load points, the effective EENS, $(EENS)_{eff}$, is given by the sum of the EENS at all the load points.

2) *Shared Maintenance*: In this maintenance strategy, the same team is assigned to perform both preventive and corrective maintenance on a component or a group of components.

3) *Dedicated Maintenance*: Unlike shared maintenance, separate teams carry out preventive and corrective maintenance on the same group of components. This implies, a failed or a component due for preventive maintenance remains unattended if its dedicated maintenance team is unavailable.

B. The Cost Model

The resultant effect of component failure, maintenance strategy and operational dynamics on the system is expressed in terms of the expected total loss, L, incurred. Assuming zero inflation, its components are expressed thus;

- Loss, L_1, due to lost output, which in turn is due to system outages, consequent of component failure and maintenance. If C_0 is the cost of a unit output, L_1 is expressed as,

$$L_1 = C_0 (EENS)_{eff} \qquad (4)$$

For commercial power systems, $EENS$ is in kWh and C_0, the cost of a kWh (e.g in $£/kWh$).

- Fixed maintenance cost, L_2, emanating from fixed wages for maintenance personnel. If each team of group j is paid r_j units of currency per unit time, L_2 is given by,

$$L_2 = T_m \sum_{j=1}^{\omega} r_j n_{t_j} \qquad (5)$$

- Total cost, L_3, associated with the fixed cost per maintenance action. This cost normally is associated with transportation of crew, contribution to offset purchasing cost of tools, or both. If m_c is the cost per maintenance action, $N_i^{\{cm\}}$ and $N_i^{\{pm\}}$ respectively the number of successful corrective and preventive maintenance actions for component i, L_3 is given by,

$$L_3 = \sum_{i=1}^{M} m_c \left(N_i^{\{cm\}} + N_i^{\{pm\}} \right)$$
$$M = \sum_{j=1}^{\omega} m_j \qquad (6)$$

where M is the number of maintainable components of the system. When expressed in closed form, (6) takes the form,

$$L_3 = \{m_c\}_{1 \times M} \{N_i^{\{cm\}}, N_i^{\{pm\}}\}_{M \times 2} \{1\}_{2 \times 1}$$
$$\mid i = 1, 2, ..., M \qquad (7)$$

- Cost, L_4, of maintaining system components; a function of the time spent by each component in maintenance and the cost per unit time of maintenance. If $C_i^{\{cm\}}$ and $C_i^{\{pm\}}$ respectively are the costs of corrective and preventive maintenance of component i per unit time,

$t_i^{\{cm\}}$ and $t_i^{\{pm\}}$, its total time spent in corrective and preventive maintenance, L_4 is expressed as,

$$L_4 = \sum_{i=1}^{M} \left(C_i^{\{cm\}} t_i^{\{cm\}} + C_i^{\{pm\}} t_i^{\{pm\}} \right) \qquad (8)$$

In closed form, (8) is given by,

$$L_4 = \{1\}_{1 \times M} \boldsymbol{l} \{1\}_{2 \times 1}$$
$$\boldsymbol{l} = \left(\{C_i^{\{cm\}}, C_i^{\{pm\}}\}_{M \times 2} \circ \{t_i^{\{cm\}}, t_i^{\{pm\}}\}_{M \times 2} \right) \qquad (9)$$

The '∘' operator denotes element-wise multiplication of two matrices.

- Cost, L_5, of spares used in maintaining system components. For most systems, on average, the spares used during preventive maintenance are minor and cheaper when compared to those used in corrective maintenance. Let $s_i^{\{cm\}}$ and $s_i^{\{pm\}}$ respectively be the number of spares used in corrective and preventive maintenance of component i. If their corresponding unit costs are respectively $C_{s_i}^{\{cm\}}$ and $C_{s_i}^{\{pm\}}$, then L_5 is expressed as,

$$L_5 = \sum_{i=1}^{M} \left(C_{s_i}^{\{cm\}} s_i^{\{cm\}} + C_{s_i}^{\{pm\}} s_i^{\{pm\}} \right) \qquad (10)$$

which in closed form condenses to,

$$L_5 = \{1\}_{1 \times M} \boldsymbol{l} \{1\}_{2 \times 1}$$
$$\boldsymbol{l} = \left(\{C_{s_i}^{\{cm\}}, C_{s_i}^{\{pm\}}\}_{M \times 2} \circ \{s_i^{\{cm\}}, s_i^{\{pm\}}\}_{M \times 2} \right) \qquad (11)$$

The overall system lost revenue, L, is given by,

$$L = \sum_{i=1}^{5} L_i \qquad (12)$$

Normally, the nominal system output and the various costs are known. Determination of L, therefore, effectively reduces to the task of estimating $(EENS)_{eff}$, $\{N_i^{\{cm\}}, N_i^{\{pm\}}\}_{M \times 2}$, $\{t_i^{\{cm\}}, t_i^{\{pm\}}\}_{M \times 2}$, and $\{s_i^{\{cm\}}, s_i^{\{pm\}}\}_{M \times 2}$ via reliability evaluation. These parameters are a function of the failure and maintenance events of the system components, and are therefore random. As a consequence, their mean/expected values are used in calculating the system lost revenue, L.

If the system reliability and performance indices, for strategy k, are represented by the function $R(\boldsymbol{n}^*, k)$, and the set of costs, by C, then, the system loss function can be expressed in the form, $L(C, R(\boldsymbol{n}^*, k))$. With $R(\boldsymbol{n}^*, k)$ known for all possible strategies, the optimal maintenance strategy can be identified and its sensitivity to variations in cost levels investigated without the need for multiple simulations.

C. Proposed Maintenance Regimes

Depending on the type of maintenance strategy in use, different system performance outcomes are possible, even with the same number of maintenance teams. For instance, in a series connected system, it may seem reasonable to postpone preventive maintenance until system failure. In such a scenario,

preventive and corrective maintenance actions are performed concurrently. Ideally, this should result in reduced system downtime and subsequent improvements in performance. This is normally the case if preventive maintenance actions are frequent, require large times, or if some components are not easily accessible, such that their maintenance inflicts significant throughput losses on the system. However, postponing a component's preventive maintenance may increase its likelihood of failure and bring with it additional costs. These costs are incurred from spares used, longer system down times, and higher maintenance intervention costs, as corrective maintenance durations normally are longer. In addition, more than one maintenance team may be required for efficient implementation of this strategy, since there may be multiple components requiring maintenance intervention when the system fails. On the downside, the teams are idle while the system is in operation but continue to receive salaries as the maintenance contract demands. A similar argument can be proffered for corrective maintenance of partially failed components, if in spite of the failure, system performance remains above a certain threshold. This procedure, however, may be counter productive if component interdependencies exist in the system, such that a degraded component affects the operation of healthy ones. Therefore, even for a system this simple, it's difficult to determine whether the procedure yields the most cost effective solution, without a detailed reliability analysis. In summary, the optimality of a given strategy depends, amongst other factors (cost levels, for instance), on the topology of the system and the non-topological functional relationships between its components.

Generally, the following regimes may be considered when deciding the promptness of preventive maintenance and major corrective maintenance of partially failed components.

1) Maintenance can be carried out at any time. The time of intervention depends only on the availability of maintenance teams.
2) Maintenance is carried out only when system output is nominal.
3) Maintenance is carried out only when a component is not in operation. This may coincide with the unavailability of the entire system or the unavailability of the subsystem to which it belongs.

When the maintenance of a component is interrupted due to delays in the availability of spares, two possible scenarios ensue.

4) The component remains shut down until spares are made available. In this case, there are no risks of incurring additional costs from failures. However, the maintenance team may be assigned to another task during the wait and there will be revenue losses as the system operates below its nominal performance level.
5) The component is put back into operation, in which case it continues to perform its normal function. This results in no loss of system output, provided it doesn't fail.

D. Solution Sequence

The regimes highlighted in Section II-C can be arranged into two groups. Regimes 1-3, define the promptness of maintenance actions and 4-5, the status of a component during maintenance interruptions. Each system component may be subjected to a combination of regimes; one from each group, giving rise to 6 possible maintenance strategies. Depending on the dynamics surrounding the operation of the system, additional strategies are applicable. For instance, on the basis of division of labour, preventive and corrective maintenance interventions could be shared or dedicated. This would lead to a total of 12 possible strategies, if considered. The corresponding component and system models are then derived for each of these strategies, in preparation for system optimization.

The optimization procedure follows a two-stage approach. In the first stage, the optimal maintenance strategy is identified by analysing each system model, with no restriction on the number of maintenance teams. For each case, the performance function, L, is determined, and the optimal strategy is identified as the one yielding the least value of L. The second stage searches for the optimal number of maintenance teams using this strategy. Here, the system is re-analysed for various values of n_{t_j}, in shared policies and various combinations of n_{1_j} and n_{2_j}, in dedicated policies. Given a component can undergo only one maintenance intervention at any instance, each n_{t_j} is bounded by $(0, m_j)$ and $\sum_{j=1}^{\omega} n_{t_j} \leq M$. In dedicated policies, both n_{1_j} and n_{2_j} are bounded by $(0, m_j)$, with the additional condition, $n_{1_j} + n_{2_j} \leq m_j$. Additional constraints may be imposed on the number of maintenance teams in each group, depending on the maintenance strategy and certain requirements set by the operator. For example, if two maintenance groups, i and j have at least one component in common, then, $n_{t_i} + n_{t_j} \leq \mid \theta_i \cup \theta_j \mid$. The operator, under economic constraints, may also impose bounds that are less than the limits already defined on the maintenance team size. Let $\boldsymbol{n}^* \mid \boldsymbol{n}^* = \{n_{t_1}, n_{t_2}, ..., n_{t_\omega}\}$ represent a combination of maintenance teams, and $\mathbb{N} \mid \mathbb{N} = \{\boldsymbol{n}_1^*, \boldsymbol{n}_2^*, ..., \boldsymbol{n}_\phi^*\}$, the set of all possible maintenance team combinations, with ϕ denoting their total. Deriving $\mathbb{N}$ entails obtaining, first, the set defined by the number of components in each group, such that, $\mathbb{N} = \{1, 2, ..., m_1\} \times \{1, 2, ..., m_2\} \times ... \times \{1, 2, ..., m_\omega\}$ and $\phi = \prod_{j=1}^{w} m_j$. Any combinations not satisfying the operator and maintenance-strategy-imposed constraints are removed.

$$(L_{max}, k_{opt}) = min\left(\{L(C, R(\infty, k))\}^{\mho}\right)$$
$$k = 1, 2, ..., \mho \quad k_{opt} \leq \mho \tag{13}$$

$$\left(L_{min}, \boldsymbol{n}_{opt}^*\right) = min\left(\{L\left(C, R\left(\boldsymbol{n}_j^*, k_{opt}\right)\right)\}^{\phi}\right)$$
$$j = 1, 2, ..., \phi \quad \boldsymbol{n}_{opt}^* \in \mathbb{N} \quad L_{min} \leq L_{max} \tag{14}$$

The optimal solution, therefore, is defined by the triplet, $\left(L_{min}, \boldsymbol{n}_{opt}^*, k_{opt}\right)$, where L_{min}, $\boldsymbol{n}_{opt}^*$, and k_{opt} are respectively the minimum system loss, the optimal maintenance team size combination, and the optimal strategy. If $R(\infty, k)$ represents the reliability/performance indices of the system under maintenance strategy k with no restrictions on the number of maintenance teams, and $\mho$, the number of strategies, (13) and (14) summarize the optimization procedure. $R(\infty, k)$

is obtained by setting the number of teams in each maintenance group to the number of components in that group. For this, components belonging to multiple groups are assumed to belong to the group with the least cost per maintenance team.

Large systems often result in a large number of candidate solutions. In such cases, it's advised to exploit smart optimization techniques like, Genetic Algorithm [3], [4], [9] and Particle Swarm optimization [5]. These, however, have not been considered, as the objective here is to provide a clear insight on the component and system modelling procedures.

III. SYSTEM RELIABILITY AND PERFORMANCE ANALYSIS

In this section, a brief description of the component and system modelling procedures is presented, with details on the algorithms invoked in the reliability evaluation process. To ensure simplicity and maintain focus on the modelling procedures, a perfect maintenance situation is assumed. It's, however, worthwhile noting that this is in no way limiting, as the framework can easily be extended to imperfect maintenance scenarios.

A. Component and System Representation

The multi-state model introduced in [22] is adopted to define the behaviour of each system component. This model takes cognisance of the various parameters required for the complete representation of a component's attributes. It accounts for the component's possible state transitions, their associated probability distributions, the performance level associated with each state, and any load restrictions imposed on the component.

The system is modelled as a graph which nodes represent the components and demand points of the system, and edges; their physical links. Defining the connectivity of the graph is a square adjacency matrix, conditioned to incorporate the efficiency of the physical links. Efficient algorithms were proposed in [22] to deduce the system flow equations from this matrix. These equations; a function of the flow properties of the components are in a format suitable for direct computation with the interior-point algorithm [29]. Given a system state vector, the actual flow through every node can be determined by updating the flow equations matrices and applying the interior-point algorithm. In addition to the advantages already outlined in Section I-B, the matrix representation of the system structure makes the procedure easily implementable on a digital computer. Readers are referred to [22] for the details on the multi-state component model and the flow equations.

B. Maintenance Modelling of Components

Consider a hypothetical series system, composed of binary-state components (components naturally existing in only 2 states) with capacity, c, equal to 1 when working, and 0, otherwise. The effects of repairs and preventive maintenance on the state-space of each system component, without maintenance delays, uncertainties, and maintenance suspensions are first presented. The resulting models are later modified and generalised for multi-state components in systems prone to maintenance delays and operational dynamics. The following
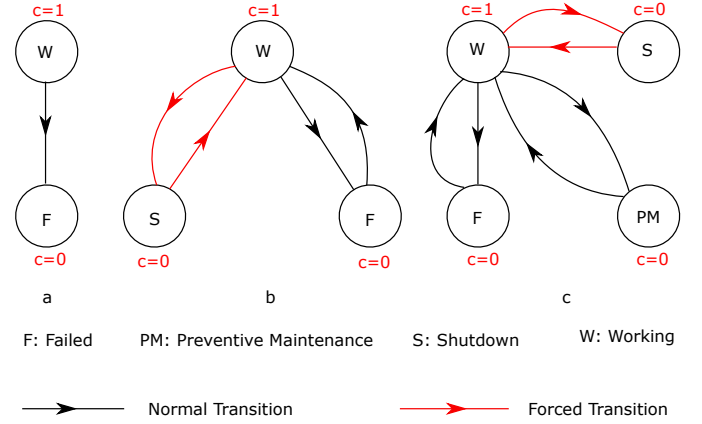

Fig. 1. State-space representation of a binary-state component under various maintenance scenarios
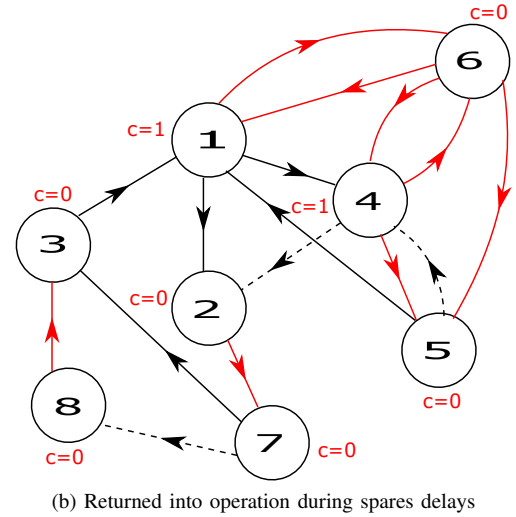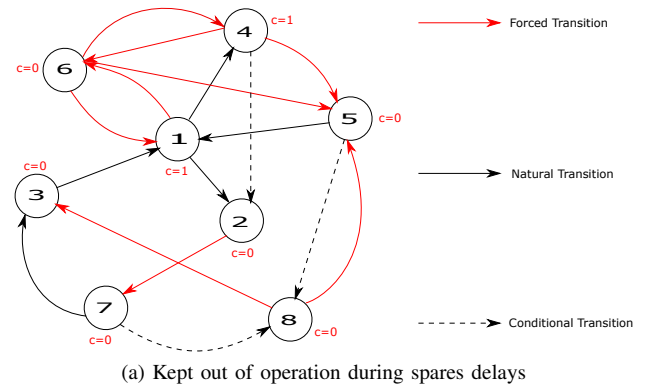

(a) Kept out of operation during spares delays

(b) Returned into operation during spares delays

Fig. 2. Repairable binary-state component under maintenance delays and operational uncertainties

maintenance scenarios are considered;

1) Each component of the system is non-repairable (Fig. 1 (a)).
2) A component can be repaired when failed (Fig. 1 (b)).
3) A component can undergo preventive, as well as corrective maintenance (Fig. 1 (c)).

Unlike the non-repairable case, a failed component is subject to repairs in scenarios 2 and 3. This is indicated by a transition from state F to state W, in Figs. 1 (b) and (c).

TABLE I
COMPONENT STATE ASSIGNMENT

State	Designation	Description
1	Working	Component operates at required capacity level.
2	Failed	Component is failed and corrective maintenance is yet to commence; $c = 0$.
3	Corrective Maintenance (CM)	Component is under repairs; $c = 0$.
4	Awaiting Preventive Maintenance (APM)	Component is due for preventive maintenance but maintenance is yet to commence; $c > 0$.
5	Preventive Maintenance (PM)	Preventive Maintenance in progress; $c = 0$.
6	Shutdown	Component not failed but taken out of operation; $c = 0$.
7	Diagnosis	Failure is being diagnosed by maintenance team; $c = 0$.
8	Idle	Diagnosis is complete but the maintenance team is waiting for spares, to resume maintenance. Required only if delays in availability of spares is modelled; $c = 0$.

TABLE II
DESCRIPTION OF STATE TRANSITIONS

Transition	Description	Transition	Description
1-2	Component Failure	7-3	Fault Diagnosis Duration
1-4	PM Interval	5-1	PM Duration
3-1	CM Duration	4-2	Failure of component whilst awaiting PM team
2-7	Forcing Diagnosis; determined by availability of maintenance team	5-8	spares needed during PM; determined by probability of spares being used
8-5	spares are available and PM resumes; determined by availability of PM team	8-3	spares are available and PM resumes; determined by availability of CM team
7-8	Spares needed during CM; determined by probability of spares being used	1-6	Shut down event like failure of system or another component
6-1,6-4	Component Restart; suggests correction of event leading to shut down	6-5	PM during shut down; determined by availability of maintenance team and whether previous state of component was APM (state 4)
4-6	Shut down event whilst component is due for PM	4-5	Forcing PM; determined by availability of maintenance team and spares
5-4	PM interruption due to spares delay		

Whilst the component is in operation, other components of the system may fail. Given a series system is unavailable with the unavailability of at least one of its components, available components are unavoidably taken out of operation during repairs of failed components. A third state, S, is therefore, introduced to account for this dependent unavailability of the operating component, as shown in Figs. 1 (b) and (c). The component remains in this state until all failed components are repaired, following which, it is restarted and restored. A fourth state, PM, is incorporated in Fig. 1 (c), to represent the period the component is in preventive maintenance.

One can easily deduce that the transitions from W to F and W to PM are competing, which is due to the perfect maintenance assumption used. Since PM and repairs make the component as good as new, any pending failures are eliminated after PM, and any scheduled PM is reset after repairs. An as good or bad as old assumption would have been implemented by replacing the transition from W to PM with a forced transition. This, however, is outside the scope of this work. It is also clear none of the three scenarios discussed considers the effects of external factors on component state transitions. For instance, there are no delays in the commencement of maintenance, and the maintenance process once initiated, suffers no obstructions or suspensions. This, however, is not
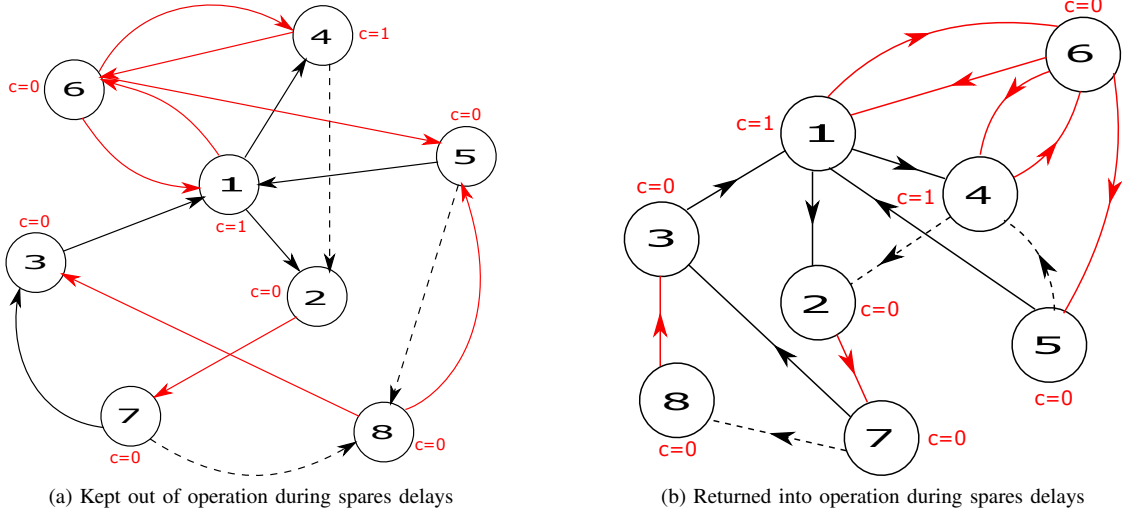
(a) Kept out of operation during spares delays

(b) Returned into operation during spares delays

Fig. 3. Repairable binary-state component under the assumption 'maintenance only when component is unavailable'

(a) Kept out of operation during spares delays
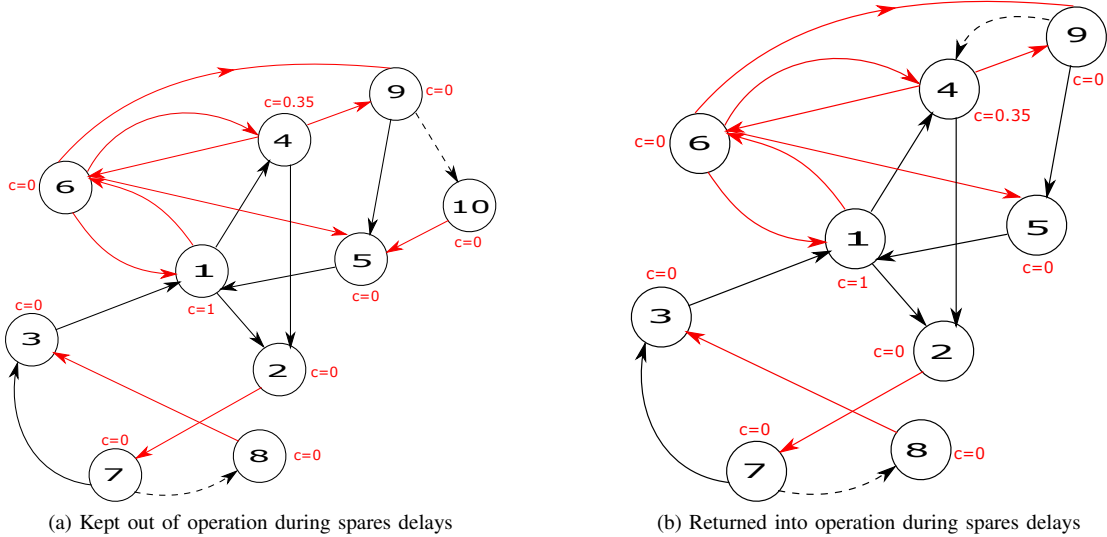
(b) Returned into operation during spares delays

Fig. 4. Repairable multi-state component under maintenance delays and operational uncertainties

the case for many practical systems.

Suppose the series system is replaced with the system described in Section II, such that there are more components than there are maintenance teams. To model such a case, four additional states are introduced in the state-space diagram in Fig. 1 (c), as shown in Fig. 2. A description of the state designations and a summary of the transitions depicted are presented in Tables I and II respectively. Fig. 2 also reveals that component state transitions can be classified as either *natural* (*normal*), *forced* or *conditional*. Natural transitions occur randomly and depend only on their associated distributions. Forced transitions occur purely as a consequence of events outside the component boundary and their distributions are unknown. Conditional transitions on the other hand, have a known distribution, but are assigned a lower priority and only occur on fulfilment of a predefined condition or a set of conditions. In the transition matrix, $\mathbf{T}_i$, of the component, conditional and and forced transitions are indicated by ∞ in

their relevant positions (see [22]). Unlike natural transitions in which the next state of a component depends only on its current state, the next state of a component under a forced transition may also depend on its previous state. For this reason, a set of special procedures are defined to execute them during system simulation.

The component models presented in Fig. 2 are based on the assumption that preventive maintenance can be carried out at any time or only when system performance is nominal. However, if preventive maintenance is carried out only when a component is out of service, the models are as presented in Fig. 3. The difference between the two sets of models is the absence of the transition from state 4 to state 5 in Fig. 3. They share the same modelling principles, as well as the designations in Tables I and II.

Multi-state component modelling under maintenance delays follows a similar approach. The models in Figs. 2 & 3 can easily be generalised for multi-state components by defining

one idle state (if components are kept out of operation during spares delay), a 'Diagnosis' state (where necessary), and one corrective maintenance state for each repairable failure mode, as shown in Fig. 4. In Fig. 4, states 4 and 5 are a partial failure mode and its corresponding corrective maintenance state respectively. States 9 and 10 are an additional 'Diagnosis' and 'Idle' states respectively, for the partial failure mode. All the other states and transitions retain their designations and meanings, as defined in Tables I and II.

C. Determining Component Transition Parameters

A system's reliability analysis by Monte Carlo simulation entails the sequential generation of the transition states and times of its components, with a view to replicating its actual operation. In a multi-state environment, a component's next transition state, y_{next}, and time, t_{next}, are determined by which of the possible transitions from its current state, x, occurs first. Given its transition matrix, all the possible transitions from state x are sampled, and the sampled times stored in a register, $\boldsymbol{Ttimes}$. The transition corresponding to the least element of this register gives the next state of the component whilst the next transition time is given by the sum of the least element and the current simulation time, t_s. In the event of multiple transitions satisfying this condition, one of them is randomly selected.

Require: x and t_s

   function SAMPLE(x)
      $J \leftarrow$ set of possible transitions from state x
      $f \leftarrow$ set of corresponding distributions
      $k \leftarrow$ Number of elements in J
      for $n \leftarrow 1$ to k **do**   ▷ Loop over possible transitions
         $(n, \boldsymbol{Ttimes}) \leftarrow$ Sample from $(n, \boldsymbol{f})$
      end for
      $t_{sample} \leftarrow min(\boldsymbol{Ttimes})$ ▷ get earliest time
      $p \leftarrow$ transitions corresponding to t_{sample}
      if $numel(\boldsymbol{p}) > 1$ **then**   ▷ if multiple transitions
         $u \sim [0,1]$ ▷ generate uniform random number
         $index \leftarrow (\lceil u * numel(\boldsymbol{p}) \rceil, \boldsymbol{p})$
      else
         $index \leftarrow \boldsymbol{p}$
      end if
      $y_{next} \leftarrow (index, \boldsymbol{J})$ ▷ get next state
      if y_{next} is APM **then**   ▷ survives till PM is due
         $t'_{sample} \leftarrow min\,(\boldsymbol{Ttimes}, t_{sample})$
         $y' \leftarrow$ state corresponding to t'_{sample}
         $t' \leftarrow t'_{sample} - t_{sample}$
      end if
      return $(y_{next}, t_{sample}, y', t')$
   end function
   $t_{next} \leftarrow t_{sample} + t_s$

Fig. 5. Algorithm 1: Sampling procedure for transition parameters of a multi-state component with preventive maintenance, under a limited maintenance team scenario

The sampling procedure described is pretty straight forward and directly applicable to most multi-state models. However, when PM is modelled as a competing transition with failures, and in the presence of limited maintenance teams, a slight modification to the procedure is required. For instance, if a working component is due for preventive maintenance (state 4 in Figs. 2 & 3) and for some reason there is a significant delay, it may fail (transition from state 4 to 2) before the commencement of maintenance. The elapsed time depends on what the failure time would be assuming the component was not subject to preventive maintenance. Therefore, if on application of the procedure, the component is found to survive till PM is due (i.e., its next state is APM), its next failure state, y', and the maximum period, t', it will survive before failure are also determined. This procedure is summarised by Algorithm 1 (Fig. 5).

1) Accounting for Non-Markovian Transitions: Algorithm 1 (Fig. 5) is only applicable to Markovian transitions (i.e., the next state of a component depends only on its current state). A second procedure, therefore, is required to implement the *forced* and *conditional* transitions. The transitions to and from shutdown, except those from shutdown to CM, PM or Diagnosis (see Figs. 2 to 4), can be implemented by the shut down and restart procedure described in [22]. The remaining conditional and forced transitions are dependent on the availability of maintenance teams or spares, where required. For these, a maintenance-forcing procedure, hinged on the assumption that the component is already assigned to an available maintenance team, is proposed.

When a component makes a transition to a new state, its next transition parameters are automatically derived, using Algorithm 1. However, for the reasons already stated, this algorithm cannot derive forced maintenance transition parameters. The component's next maintenance state, y_m, from the new state, is therefore, manually determined from its transition matrix. With correct modelling according to the models proposed in Section III-B, each failure mode will have at most one maintenance state (CM or Diagnosis) associated with it. The component is added to the corrective maintenance queue, if y_m exists. If on the other hand, the new state is APM, the transition parameters of the component are not obtained by another application of Algorithm 1. They are determined from y' and t'; obtained when the algorithm was applied when the component entered the Working state (state W). In this case, y_m is the only PM state, and the component is added to the preventive maintenance queue.

In the most general case, y_m could either be Diagnosis, CM or PM. To force maintenance, y_m is made the current state of the component, and Algorithm 1 is applied to determine its next transition parameters. It's deducible from the component models presented in Figs. 2 to 4 that a component in Diagnosis (state 7) can either undergo a normal transition to CM (state 3) or a conditional transition to Idle state (state 8). However, the sampling algorithm always yields the normal transition. Given the conditional transition to Idle state occurs only if spares are required, a uniform random number, u, between 0 and 1 is generated and compared to the probability, p_i, of spares being needed to complete the maintenance. The Idle state (state 8) is made the next transition state if $u \leq p_i$, and the transition time yielded by the sampling algorithm is retained. In the case

Require: p_i, q_i, k_i, $s_i^{\{cm\}}$, $s_i^{\{pm\}}$, t_s, y_m, $\mu_i^{\{cm\}}$, $\mu_i^{\{pm\}}$

1: **function** FORCEMAINTENANCE(i, $input$)
2:     $x \leftarrow y_m$ ▷ Force transition
3:     $(y_{next}, t_{sample}, \sim, \sim) \leftarrow$ SAMPLE(x)
4:     **if** x is PM **then**   ▷ In preventive maintenance
5:         **if** $\mu_i^{\{pm\}} \leftarrow 1$ **then**   ▷ From suspension
6:             $t_{sample} \leftarrow (\frac{1}{k_i} - 1)t_{spent}$
7:             $\mu_i^{\{pm\}} \leftarrow 0$ ▷ Reset indicator
8:         **else if** $u \sim [0, 1] \leq q_i$ **then**   ▷ Spares needed?
9:             $s_i^{\{pm\}} \leftarrow s_i^{\{pm\}} + 1$ ▷ add PM spares used
10:             $t_{sample} \leftarrow k_i t_{sample}$
11:             $x_{prev} \leftarrow$ previous state
12:             **if** $\mathbf{T}_i(x, x_{prev}) \neq 0$ **then**   ▷ If to restart
13:                 $y_{next} \leftarrow x_{prev}$
14:             **else**
15:                 $y_{next} \leftarrow$ 'Idle' state linked to x
16:             **end if**
17:             $\mu_i^{\{pm\}} \leftarrow 1$ ▷ Set indicator
18:         **end if**
19:     **else if** x is Diagnosis **then**
20:         **if** $\mu_i^{\{cm\}} \leftarrow 1$ **then**   ▷ From suspension
21:             $x \leftarrow$ CM state connected to x
22:             $\mu_i^{\{cm\}} \leftarrow 0$ ▷ Reset indicator
23:             $(y_{next}, t_{sample}, \sim, \sim) \leftarrow$ SAMPLE(x)
24:         **else if** $u \sim [0, 1] \leq p_i$ **then**
25:             $s_i^{\{cm\}} \leftarrow s_i^{\{cm\}} + 1$ ▷ add CM spares used
26:             *call lines 12 to 16*
27:             $\mu_i^{\{cm\}} \leftarrow 1$ ▷ Set indicator
28:         **end if**
29:     **end if**
30:     $t_{next} \leftarrow t_{sample} + t_s$
31:     **return** $(y_{next}, t_{next}, s_i^{\{cm\}}, s_i^{\{pm\}}, \mu_i^{\{cm\}}, \mu_i^{\{pm\}})$
32: **end function**

Fig. 6. Algorithm 2: Procedure for forcing maintenance

of repair from a partial failure mode, such that the component is returned into operation during spares delay (see states 4 and 9 in Fig. 4 (b)), the partial failure mode is made the next state, and $\mu_i^{\{cm\}}$, assigned the value 1. $\mu_i^{\{cm\}}$ is an indicator function that takes the value 1 when CM is suspended, and 0, otherwise. The component is removed from the maintenance queue until the spares requested are made available.

$$t_{spent} = k_i t_{pm}$$
$$t_{next} = t_s + (1 - k_i)\, t_{pm}$$
$$= t_s + \left(\frac{1}{k_i} - 1 \right) t_{spent} \tag{15}$$

Similarly, a component in preventive maintenance (state 5 in Figs. 2 and 3) can either return to Working state (state 1), go to Idle state (state 8), or return to its previous state, if it should be kept in operation whilst awaiting spares. Like corrective maintenance, any of the last two outcomes is determined by the probability, q_i, of spares being needed to complete preventive maintenance. The next transition time if spares are

required is given by $t_s + k_i t_{pm}$, where t_{pm} is the PM duration yielded by Algorithm 1, and k_i, its proportion spent before the maintenance team realises spares are required. When PM is suspended, the component is removed from the maintenance queue, and $\mu_i^{\{pm\}}$, its indicator function for PM suspension, set to value 1. On PM resumption, the expected duration of the remainder of the maintenance exercise is $(1 - k_i)\, t_{pm}$. To avoid storing too many variables during simulation, this period is expressed in terms of t_{spent}, the time spent by the component in PM before maintenance suspension. t_{spent} is computed from the saved transition history of the component, and the next transition time, t_{next}, is derived as in (15). The maintenance-forcing procedure described above is summarised by Algorithm 2 (Fig. 6).

D. Maintenance Strategy Implementation

Algorithm 2 assumes the component has already been assigned an available maintenance team. However, with multiple components requiring maintenance intervention, maintenance team assignment follows the maintenance strategy in use. Let h_1 and h_2 respectively be the sets of components requiring CM and PM, $\Pi = \{n_{1_j}, n_{2_j}\}_{\omega \times 2} \mid j = 1, 2, ..., \omega$, the matrix defining the number of CM and PM teams in each maintenance group, and $\varphi = \{\varphi_j\}_{\omega \times 1}$, an indicator vector which elements are matched to the rows of Π.

$$\varphi_j = \begin{cases} 1, & \text{If maintenance group j is shared} \\ 0, & \text{Otherwise} \end{cases} \tag{16}$$

Each indicator element specifies whether its corresponding maintenance group practices shared or dedicated maintenance, as defined by (16).

Given the assumption of a component being as good as new after PM or CM and the additional constraint that the former is carried out only on the perfect component, the condition $h_1 \cap h_2 = \emptyset$ is imposed. Therefore, prior to maintenance team assignment, all the elements of $h_1 \cap h_2$ are removed from h_2 (i.e., $h_2 = h_2 - (h_1 \cap h_2)$ or simply $h_2 = h_2 - h_1$). Depending on the maintenance strategy, additional components may be removed from h_1 and h_2. For instance, if Ω is the set of components in Shutdown state, η_1, the set of components repairable only while in Shutdown state, and η_2, the set of components which PM is initiated only when in Shutdown, then, $h_1 = (h_1 - \eta_1) \cup (\Omega \cap \eta_1)$ and $h_2 = (h_2 - \eta_2) \cup (\Omega \cap \eta_2)$. Similarly, let δ_1 be the set of components repairable only while system performance is nominal, and δ_2, the set for which PM is initiated only at nominal system performance. If system performance is below nominal at maintenance team assignment, $h_1 = h_1 - \delta_1$ and $h_2 = h_2 - \delta_2$. Please note η_1 applies to partially failed components only.

With h_{1f} and h_{2f} representing the final contents of h_1 and h_2 respectively, the first maintenance group is considered. Its assigned components in the maintenance queue are ranked according to the predefined priority rule and the top ranked component is assigned to the first available team in the group. As a consequence, the number of available teams and the number of ranked components reduce by 1 each. The

Require: $(h_{1f} \cup h_{2f}) \neq \emptyset$, h_1, h_2

1: **for** $j \leftarrow 1$ to ω **do** ▷ Loop over maintenance groups
2:    **if** $\varphi_j > 0$ **then** ▷ If maintenance is shared
3:       $Teams \leftarrow \mathbf{\Pi}\,(j,1) + \mathbf{\Pi}\,(j,2) - \left(\lambda_j^{\{cm\}} + \lambda_j^{\{pm\}}\right)$
4:       $X_{comp} \leftarrow \left(h_{1f} \cap \theta_j^{\{cm\}}\right) \cup \left(h_{2f} \cap \theta_j^{\{pm\}}\right)$
5:       **while** $Teams > 0$ **and** $X_{comp} \neq \emptyset$ **do**
6:          $i \leftarrow$top ranked component
7:          FORCEMAINTENANCE(i, $input$)
8:          $Teams \leftarrow Teams - 1$
9:          $X_{comp} \leftarrow X_{comp} - i$ ▷ remove component
10:          *adjust X_{comp} if necessary*
11:       **end while**
12:    **else**
13:       $\mathbf{H} \leftarrow \{h_{1f}, h_{2f}\}$
14:       $\mathbf{G} \leftarrow \{\theta_j^{\{cm\}}, \theta_j^{\{pm\}}\}$
15:       $\mathbf{I} \leftarrow \{\lambda_j^{\{cm\}}, \lambda_j^{\{pm\}}\}$
16:       **for** $k \leftarrow 1$ to 2 **do**
17:          $X_{comp} \leftarrow (k, \mathbf{H}) \cap (k, \mathbf{G})$
18:          $Teams \leftarrow \mathbf{\Pi}(j, k) - (k, \mathbf{I})$
19:          *call lines 5 to 11*
20:       **end for**
21:    **end if**
22:    *Remove assigned components from h_{1f} and h_{2f}*
23:    **if** $(h_{1f} \cup h_{2f}) \leftarrow \emptyset$ **then**
24:       **break**
25:    **end if**
26: **end for**
27: *Remove assigned components from h_1 and h_2*

Fig. 7. Algorithm 3: Procedure for maintenance strategy implementation during simulation

procedure is repeated until all the ranked components have been assigned or until there are no available maintenance teams in the group. At this stage, h_{1f} and h_{2f} are updated accordingly, and the next maintenance group considered if $h_{1f} \cup h_{2f} \neq \emptyset$. This recursive procedure continues until all the maintenance groups have been covered.

Let $\theta_j^{\{cm\}}$ be the set of components assigned to maintenance group j for CM and $\theta_j^{\{pm\}}$, the set assigned for PM. If $\lambda_j^{\{cm\}}$ and $\lambda_j^{\{pm\}}$ are respectively the numbers of unavailable teams from group j for CM and PM, Algorithm 3 (Fig. 7) summarises the maintenance strategy implementation procedure. Line 10 accounts for the case when components maintained only while system performance is nominal are removed from the queue following the deviation from nominal performance. This normally is a consequence of either PM or CM of a partially failed component of a higher rank in the queue.

E. The Simulation Procedure

A discrete-event simulation model is proposed to replicate the behaviour of the system. Starting with components in their initial states, the initial performance level of the system is computed and recorded. Following which, the next transition parameters of each component are sampled, and the simulation progresses to the earliest transition time. At this time, the current state of the appropriate component making the transition is updated, its new state recorded as a function of time, its next transition parameters sampled, and the next simulation time determined. This procedure is repeated for subsequent transitions until the mission time is exceeded. For every transition resulting in a change in the flow properties of a component, the output of the system is computed and recorded as a function of time. The relevant reliability and performance indices are determined from the saved component transition and system output histories.

Let $\boldsymbol{\tau}$ be the vector of next transition times of nodes (components and output points) and $\boldsymbol{\tau}_{spare}$, the vector holding the availability times of component spares. If M' is the total number of system nodes, the simulation procedure is summarised as follows,

1) Initialise the system in preparation for simulation. This involves,
   a) initialization of registers to save the current flow properties of nodes, transition history of components, and the performance histories of output nodes.
   b) setting the required number of simulations, $Nsamples$, and mission time, T_m.
2) Set $t_s = 0$, $s_i^{\{cm\}} = s_i^{\{pm\}} = \mu_i^{\{cm\}} = \mu_i^{\{pm\}} = 0 \forall i \in \{1, 2, ..., M'\}$, $h_1 = h_2 = \emptyset$, $\boldsymbol{\tau} = \boldsymbol{\tau}_{spare} = \{\infty\}^{M'}$.
3) Save the initial states of components.
4) Compute the initial performance level of all the output nodes, and save as a function of t_s.
5) Sample the next transition parameters (y_{next} and t_{next}) of nodes, update $\boldsymbol{\tau}$ and set $t_s = min\,(\boldsymbol{\tau})$.
6) Check for nodes with next transition time equal to t_s. For each node, i,
   a) effect the required transition.
   b) with the exception of the case when the new state is APM, Idle, or partial failure given its previous state is Diagnosis, sample its next transition parameters and determine y_m, where applicable. Update h_1 or h_2 if y_m exists, set $\mu_i^{\{cm\}}$ and $\mu_i^{\{pm\}}$ to 0, and go to Step (g).
   c) if the new state is APM, $y_{next} = y'$, $t_{next} = t' + t_s$, y_m is set to the PM state, and h_2 updated. However, h_2 is not updated if the node is returning from PM, as the transition depicts a maintenance suspension. In this case, $t_{next} = t_{rem} + t_s$, where t_{rem} is the remaining life of the component prior to its maintenance being forced. Go to Step (f).
   d) if the new state is partial failure and previous state Diagnosis, $t_{next} = t_{rem} + t_s$, the expected failure state before the transition to Diagnosis is made y_{next}, and y_m is set to Diagnosis. Go to Step (f).
   e) if the new state is Idle, $t_{next} = \infty$. y_m is set to PM if the node is from PM, and CM if it is from Diagnosis. Go to Step (f).
   f) Steps (d) and (e) involve maintenance suspensions. For these and the case involving PM suspension in

TABLE III
COMPONENT AND SYSTEM DATA FOR THE HYDRO ELECTRIC POWER PLANT

Component	Valves	Turbines	Gens.	Breakers	Synch.	Xfmr.
Failure time distribution	Wb(1000,1.5)	Wb(4125,2.1)	Wb(2000,2)	Exp(3750)	Exp(3250)	Exp(2500)
Repair time distribution	Exp(40)	LogN(106,5)	Exp(150)	Exp(36)	Exp(96)	Exp(80)
PM interval	U(500,625)	U(1125,1250)	U(1125,1250)	U(2125,2175)	U(2125,2175)	U(2125,2175)
PM duration	Exp(8)	Exp(21.2)	Exp(30)	Exp(7.2)	Exp(19.2)	Exp(16)
Diagnosis duration	Exp(5)	Exp(14)	Gu(20,3.24)	G(5,2)	Exp(16)	LogN(16,2)
Spares cost(CM)	1624	2100	1944	1006	2245	2700
Spares cost(PM)	1055.6	1365	1263.6	653.9	1459.25	1755
PM cost/hr	162.5	243.75	203.13	101.56	243.75	264.06
CM cost/hr	250	375	312.5	156.25	375	406.25
Spares delay	Exp (24)					
Probability of Component Replacement During Maintenance						
CM (p_i)	0.5	0.55	0.8	0.9	0.7	0.6
PM (q_i)	0.8	0.9	0.96	0.42	0.4	0.45
Mean Fraction of PM Duration Before Component Replacement Becomes Eminent						
Fraction (k_i)	0.25	0.25	0.25	0.25	0.25	0.25

Step (c), the time, t_{spare}, the spares will be delayed by is sampled from the appropriate distribution. Update $\boldsymbol{\tau}_{spare}$, such that $(i, \boldsymbol{\tau}_{spare}) = t_{spare} + t_s$.

   g) Update the node's state history, the flow property vectors, and $\boldsymbol{\tau}$, such that $(i, \boldsymbol{\tau}) = t_{next}$.

7) Identify nodes for which spares have been made available, that is, $(i, \boldsymbol{\tau}_{spare}) = t_s$. For each node, i, update $\boldsymbol{\tau}_{spare}$, such that $(i, \boldsymbol{\tau}_{spare}) = \infty$, h_1 if y_m is CM or Diagnosis, and h_2, otherwise.

8) Compute h_{1f}, h_{2f} and call Algorithm 3 (Fig. 7).

9) If the current and previous flow property vectors differ,
   a) restart nodes in shut down, compute system flow and shut down nodes, as proposed in [22].
   b) for each output node, update its performance history if its current and previous performances differ.

10) Save the current node flow property vectors.

11) Compute $h_{1f} = h1 \cap \Omega \cap \eta_1$, $h_{2f} = h2 \cap \Omega \cap \eta_2$ and call Algorithm 3 for the second time. This step accounts for those components maintainable only while in Shutdown.

12) Set the next simulation time, $t_s = min(min(\boldsymbol{\tau}), min(\boldsymbol{\tau}_{spare}))$.

13) Repeat Steps 6 to 12 until $t_s > T_m$, updating $\boldsymbol{\tau}$, the flow property vectors, node state histories, and output performance histories at every transition.

14) Repeat Steps 2 to 13, $Nsamples$ times, saving the final node histories at every trial.

15) Determine the system performance indices.

The desired performance indices are, $(EENS)_{eff}$, $\{N_i^{\{cm\}}, N_i^{\{pm\}}\}_{M \times 2}$, $\{t_i^{\{cm\}}, t_i^{\{pm\}}\}_{M \times 2}$, and $\{s_i^{\{cm\}}, s_i^{\{pm\}}\}_{M \times 2}$. The latter is yielded directly by the simulation algorithm, $(EENS)_{eff}$ is computed from the performance histories of output nodes, and the remainder, from the state transition histories of components. $t_i^{\{pm\}}$ is given by the average time spent by component i in PM state (e.g., state 5 in Figs. 2 and 3), $t_i^{\{cm\}}$, the average time spent in Diagnosis and CM (e.g., states 7 and 3 in Figs. 2 and 3, states 3, 5, 7 and 9 in Fig. 4), $N_i^{\{cm\}}$, the average number of transitions from all CM states to Working state (e.g., transition 3-1 in Figs. 2 and 3, transitions 3-1 and 5-1 in Fig. 4) and $N_i^{\{pm\}}$, the average number of transitions from PM

state to Working state (e.g., transition 5-1 in Figs. 2 and 3). These indices are substituted in the equations proposed in Section II-B, to compute the system loss function.

The simulation procedure, with its associated algorithms, accounts for most of the forced and conditional transitions. As a result, an appreciable number of these transitions could be omitted from the component model with no adverse effects on the simulation outcome. For instance, the Shutdown state and its related transitions could be omitted altogether. This, however, does not mean shut down and restart are not accounted for during simulation. Of the remaining forced and conditional transitions, only those to and from Diagnosis state, from PM to Idle state, and from PM to APM state (if applicable) are required, the rest could be omitted. Applying this new information to the component models presented in Figs. 2 to 4, for instance, would result in much simpler models.

IV. CASE-STUDIES

The proposed framework is implemented in the open source MATLAB-based toolbox, OpenCOSSAN [30], [31], and used to identify the optimal maintenance strategies for two power systems.

A. Case-Study 1: A 50MW Hydroelectric Power Plant

In this case-study, a two-unit hydroelectric power plant is analysed. It is a slightly modified model of the Bumbuna hydroelectric power plant; a 50MW plant in Sierra Leone. Its two units are similar, and each, rated 25MW consists a butterfly valve, turbine, generator, and circuit breaker. Their generated power is synchronized in the synchronizing unit and fed to the step-up transformers for onward transmission. These transformers are also rated 25MW, and when one is unavailable, the plant is reconfigured such that only one unit operates. The plant's schematic representation is shown in Fig. 8 and its reliability data, in Table III. All failure and repair times are in hours, and costs, in British Pounds (£). The unit cost of electricity is £ 0.5, the fixed wage per maintenance team is £ 7 per hour, and a negligible cost per maintenance call. It is worthwhile noting that the data presented in Table III are assumed, and therefore for illustrative purposes only.
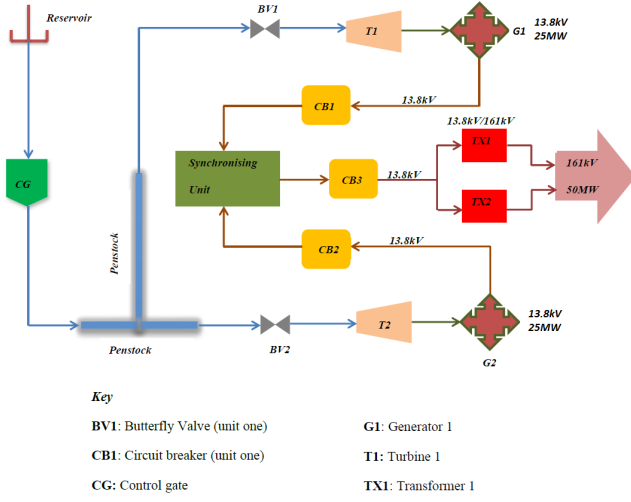
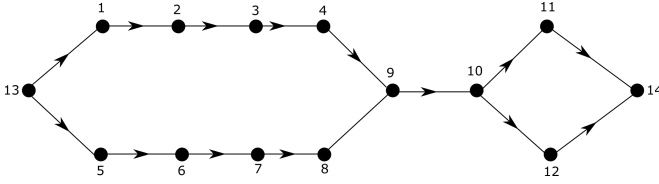Fig. 8. Schematic of a 2-unit hydro electric power plant
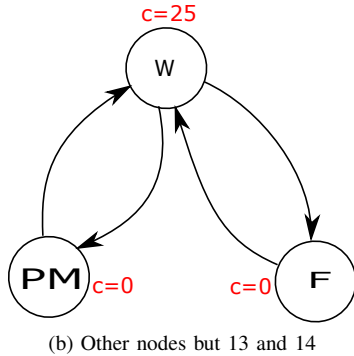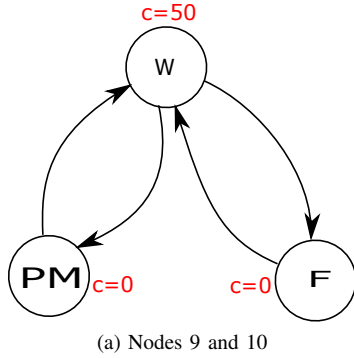

Fig. 9. Plant's network model

(a) Nodes 9 and 10

(b) Other nodes but 13 and 14

Fig. 10. State Space diagrams of components

Fig. 11. Plant output performance

Fig. 12. Plant reliability

Ideally, such data are based on actual field data extracted from component maintenance history.

1) Modelling the Plant and its Components: The following assumptions are considered;
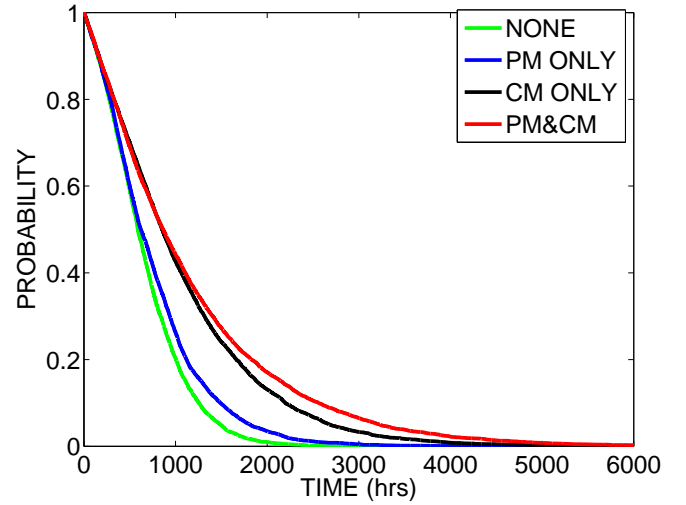
1) All components operate at only two distinct performance levels.
2) Components are ranked for maintenance in their order of arrival in the maintenance queue.
3) There is only one maintenance group.
4) The load on the plant is fixed at 50MW, and there is sufficient water in the reservoir to meet this demand.
5) The failure rates of the control gate and penstock are negligible.

Fig. 9 shows the network model of the plant. The components of unit 1; valve-1, turbine 1, generator-1 & breaker-1 are respectively represented by nodes 1 to 4 and their counterpart in unit 2, by nodes 5 to 8. Nodes 9 to 14 respectively represent the synchronizer, breaker-3, transformer-1, transformer-2, dam, and the external load. Assuming perfect links between components, the parameters of the network are obtained as proposed in [22]. For this system, the number of nodes, M', is 14 while the number of maintainable components, M, is 12. The state-space diagrams of the components, without maintenance delays are shown in Fig. 10. Under the range of

possible maintenance regimes proposed in Section II-C, these state space diagrams can be transformed into those in Figs. 2 and 3. Since the demand and source (dam) capacity are fixed at 50MW, nodes 13 and 14 have a single state of capacity 50 units.

The reconfiguration procedure used in the simulation shuts down nodes when their load flow drops below a threshold. To enable plant reconfiguration when only one transformer is available, a minimum load restriction is imposed on the turbines. The choice of the turbines, however, is arbitrary, as any of the unit nodes would do, due to their being connected in series. With only node 11 or 12 available, the load flow from node 13 drops to 25MW, which is divided equally between the two units if they both are in operation. The threshold flow for each turbine, therefore, is set to a value slightly greater than 12.5 units (say 12.52), and 0, for all the other nodes.

2) The Effects of Maintenance on System Performance and Reliability: The plant is analysed separately under the assumptions that its components are non-repairable, subject to PM only, CM only and both maintenance types. With the exception of the non-repairable case, there is no restriction on the number of parallel maintenance actions that can take place. The maintenance team size in each case, therefore, is expressed as (0 0), (0 12), (12 0) and (12 0) respectively. Dedicated maintenance is used in the second and third cases to ensure only the intended maintenance type is carried out (e.g., no CM during a PM only policy). This stage of the optimization is aimed at investigating the relative effects of the various maintenance strategies on the plant's reliability, performance and loss function. It identifies the candidates for the optimal maintenance strategy and determines whether or not to proceed with the search for the optimal maintenance team size. This prevents searching in unlikely regions or strategies, thereby reducing the computational cost.

Figs. 11 and 12 respectively show the reliability and instantaneous performance of the plant as a fraction of its nominal output, for a mission time of 10^4 hours and 5×10^3 Monte Carlo samples. Plant reliability is defined with respect to complete outages, however, excluding those due to preventive maintenance (scheduled outages). The objective is to study the survivability of the plant, which scheduled outages would underestimate. For instance, more frequent outages may be experienced under a maintenance strategy incorporating both PM and CM than one with CM only. In practice, scheduled outages do not count toward a systems's survivability, since they are out of choice rather than failure. Hence, the need for their disregard in its survivability analysis. In summary, plant reliability at time, t, is the non-occurrence probability of complete-outage-inducing failures in the interval $[0, t]$.

The reliabilities and instantaneous performances defined by Figs. 11 and 12 depict the upper bounds for the various maintenance strategies. As expected, both types of maintenance (PM & CM) action, indeed improve the reliability and performance of the plant. The impact of PM, however, is only slight, given that 50% of the components exhibit an exponential failure characteristic. For such components, PM only reduces their availability without an improvement in reliability [23]. Preventive maintenance, therefore, is most effective in systems

TABLE IV
PLANT EXPECTED OUTPUT AND LOSS

Strategy		Output (GWh)	L ($£10^6$)
None		23.6646	238.17
PM only		26.0639	237.82
CM only		382.2114	60.98
PM+CM	[1,4]	370.9891	66.38
	[1,5]	384.2075	59.91
	[2,4]	369.1798	67.51
	[2,5]	383.5723	61.42
	[3,4]	396.2899	53.63
	[3,5]	388.2218	58.07

TABLE V
OPTIMAL PLANT LOSS AS A FUNCTION OF MAINTENANCE STRATEGY

Strategy	L ($£10^6$)	Number of teams
[1,4]	65.6617	2
[1,5]	59.2353	2
[2,4]	66.8779	3
[2,5]	59.6466	3
[3,4]	52.8917	5
[3,5]	57.3184	4
CM only	60.1399	4

with ageing components. Table IV presents the upper bound of the expected plant output and the corresponding loss for each maintenance strategy. The notation [a,b] denotes a strategy made up of a combination of regimes *a* and *b*, as described in Section II-C. A review of the trend portrayed in Figs. 11, 12, and Table IV suggests a maintenance strategy incorporating both PM and CM is desirable. The losses in Table IV are yielded by the maximum number of maintenance teams, the optimal loss in each case, therefore, will be provided by fewer maintenance teams. These teams are determined by the procedure proposed in Section II-D.

3) Optimal Maintenance Strategy Identification: It's clear the non-repairable and 'PM only' strategies are very inefficient. The plant, therefore, is analysed for the other strategies, using the same mission time and number of samples as in the preceding section. The optimal solution for each strategy is identified and recorded as shown in Table V. From these, the best maintenance strategy and the optimal number of maintenance teams are deduced as [3,4] and 5 respectively. To explore the existence of a more optimal solution for this strategy, the plant is re-analysed under dedicated maintenance. It's observed that for the same number of teams, shared maintenance strategies produce a better plant performance.

The optimal strategy being [3,4] is in agreement with the preliminary results presented in Table IV. Therefore, the optimal solution would have been obtained using this strategy alone. However, the other strategies were considered to establish a relationship between the optimal maintenance team size and maintenance strategy.

4) Sensitivity to Cost Levels: The robustness of the optimal maintenance strategy to variations in cost of electricity (EC), fixed cost per maintenance team (FMC), fixed cost per hour of maintenance (CPHM), and cost of spares (CS) is investigated. Fig. 13 shows how the number of maintenance teams required for optimal performance varies with $k_f \mid 0 \leq k_f \leq 100$, where k_f is the ratio of new cost to the original cost provided in Table

TABLE VI
OPTIMAL MAINTENANCE STRATEGY SENSITIVITY TO COSTS

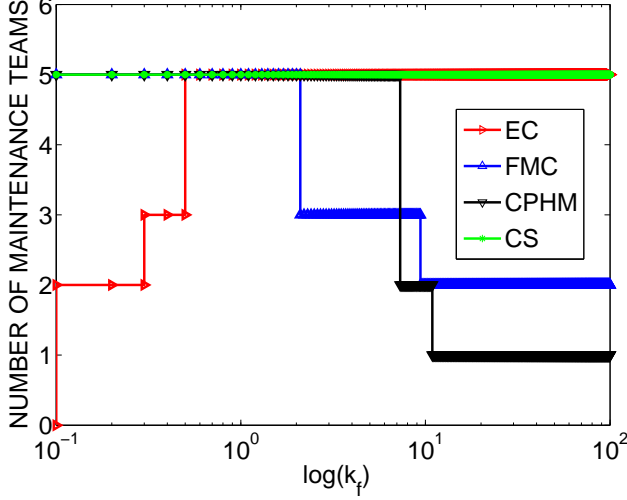| | Cost Element | | | |
	EC	FMC	CPHM	CS
Strategy	$(0\ 0), k_f = 0$ $[3,4]\quad k_f > 0$	$[3,4]\quad \forall k_f$	$[3,4], 0 \le k_f < 70.9$ $[3,5],\qquad k_f \ge 70.9$	$[3,4]\quad \forall k_f$

Fig. 13. Optimal maintenance team size sensitivity to costs
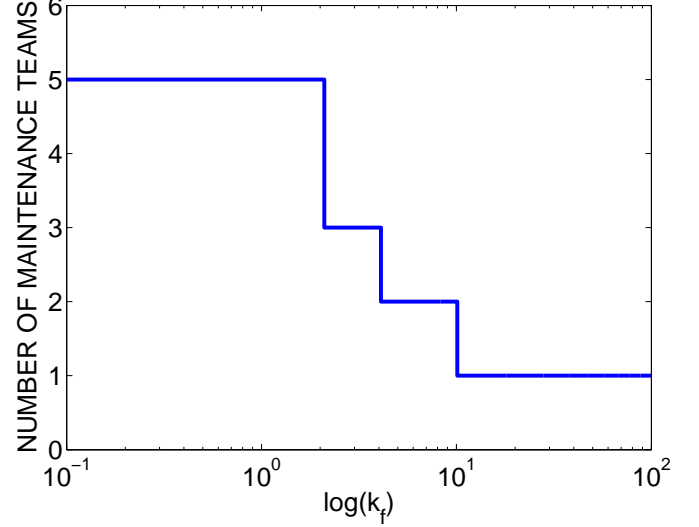
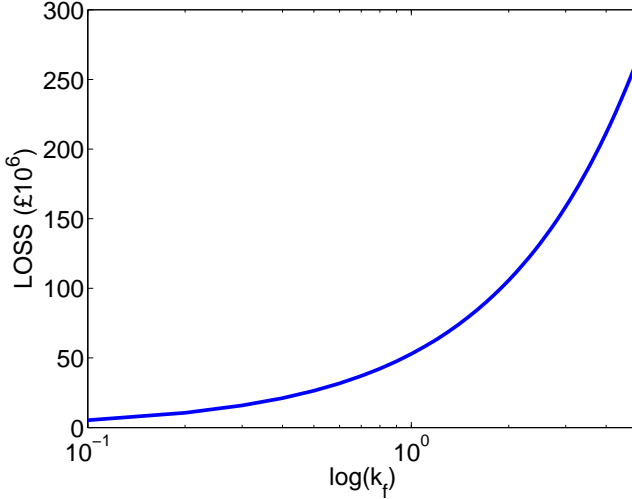Fig. 15. Sensitivity of optimal solution to concurrent variation in FMC and CPHM

Fig. 14. Optimal system loss sensitivity to cost level variation

III. It's evident from the figure that the optimal maintenance team size is insensitive to the cost of spares but exhibits a fair degree of sensitivity to the other costs. In contrast, the optimal maintenance strategy is insensitive to all four cost elements up to $k_f = 70.9$ (for CPHM), beyond which [3,5] becomes the optimal strategy, as shown in Table VI.

In practice, when inflation occurs, it affects all the cost elements concurrently. The sensitivity of the optimal solution in such a scenario is investigated. It is observed that with $k_f = 0$, the maintenance strategies are all equivalent, since all the services are basically provided free-of-charge. Beyond this value, the optimal maintenance strategy and the number of

teams remain constant at [3,4] and 5 respectively, for the entire range of k_f. The optimal loss, however, increases according to Fig. 14. This strange behaviour is explained by the dominance of the cost of electricity in the loss equation (see Section II-B). When all the four costs change by the same factor, the resultant effect is dominated by the electricity cost, for $k_f > 0.4$, and the other costs, otherwise. A comparison of the trends portrayed in Figs. 13 and 15 supports this theory. Fig. 15 is obtained by holding fixed, the cost of electricity and varying the maintenance costs. Expectedly, it shows a decrease in the optimal maintenance team size, with rising maintenance costs. Indeed, with high maintenance costs, the only logical decision is downsizing the maintenance team to ensure sustainability.

5) Computational costs: The simulations were run on a *48 core, 1895.257MHz AMD Opteron(tm) 6168 processor* using 19 cores running in parallel. Less than one minute was required for the non repairable system and an average of 8.95 minutes per candidate solution for the system under preventive and corrective maintenance.

6) Discussions: Analytical approaches do not make a feasible option for the analysis of complex systems with realistic attributes. Simulation algorithms on the other hand are disadvantaged by their large computational costs, made worse when employed in optimization procedures. This, often, is attributed to the large number of samples required for a dependable estimate of the system performance indices. Therefore, the trade-off between accuracy and moderate computational burden is worth adequate attention. Another limiting constraint of great importance is the mission time, which should be selected such

that, the performance indices obtained reflect the true long-term indices of the system. This requires that the mission time be sufficiently greater than the time the system takes to attain steady-state. In the case-study presented, 5000 samples are just enough to provide an acceptable degree of accuracy and a manageable computational burden. Also, as deduced from Fig. 11, the plant's steady-state attainment time is about a fifth of its mission time. These attributes endorse the dependability of the optimization outcome.

The analyses suggest the optimal number of maintenance teams is maintenance strategy-dependent. They also reveal, returning components into operation during maintenance suspensions improves system performance. This improvement is attributable to the increased availability of the components culminating in a lower EENS. The exception is the case when PM is initiated only while components are not in operation. In this regime, the initiation of a component's PM is determined by the failure characteristics of other components. Therefore, when the component is returned into operation, its PM resumes only on the occurrence of another shut down event. The likelihood that the component fails in this interval is higher than in the other regimes, due to the longer wait times. The result is, a fewer PM actions, more failures, longer component downtimes, and a higher EENS. These consequences are minimized by keeping the component out of operation until PM resumes. However, in both cases, initiating PM only while components are not in operation yields the best performance.

The range of k_f used in the sensitivity analysis is a little unrealistic for practical applications. The range of interest, therefore, is conservatively chosen to be $0 \leq k_f \leq 2$, depicting an inflation of -100% to $+100\%$. In this range, the optimal maintenance strategy is unaffected by variations in cost levels, though the number of teams required for optimal performance varies with the cost of electricity. The following, therefore, is recommended for the hydro electric power plant.

1) Preventive maintenance should be carried out only when a component is not in operation. That is, it should coincide with a shut down event that renders the component inactive.
2) Components should be kept out of operation during maintenance interruptions.
3) At the current cost levels, 5 maintenance teams; in a shared maintenance strategy are required for optimal performance. However, this should be scaled down to 3, 2, 1 and 0 when the cost of electricity deflates by 50%, 60%, 90% and 100% respectively (see Fig. 13).
4) As evidenced in Figs. 11 and 12, preventive maintenance does not quite improve the overall performance of the system, contrary to anticipations. This, as explained earlier, could be due to subjecting components exhibiting exponential failure characteristics to needless preventive maintenance. It's anticipated that if preventive maintenance is not carried out on these components, additional gains could be made from improved plant availability and reduced maintenance costs. This hypothesis is tested, and as expected, results in an output gain of 1.82% and a corresponding system loss reduction by 7%. Preventive maintenance, therefore, should not be carried
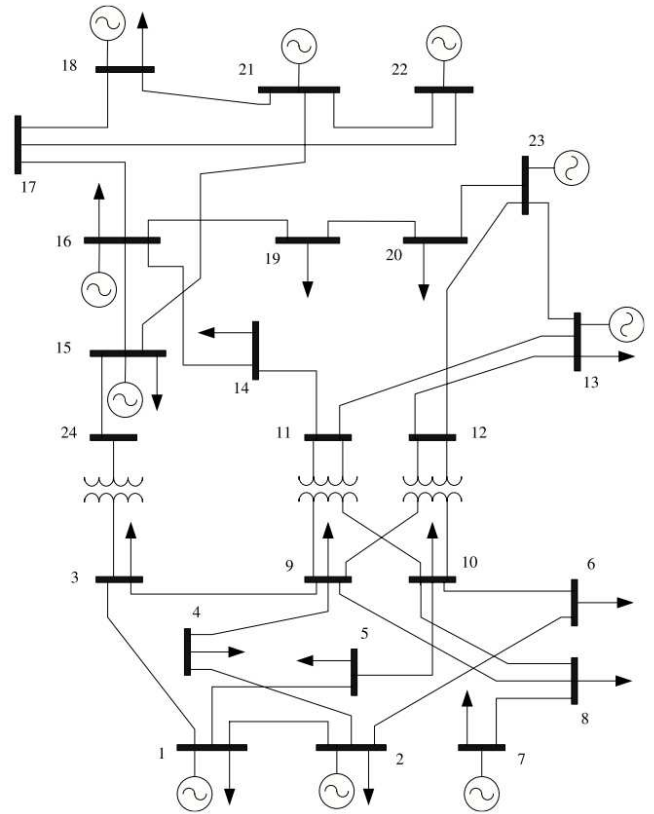

Fig. 16. Single line diagram of the IEEE-24 bus Reliability Test System

out on the breakers, synchronizer and transformers.

B. Case-Study 2: The IEEE-24 Bus Reliability Test System

In this case-study, we consider a more realistic system, in order to showcase the applicability of the proposed approach to systems of practical nature. Shown in Fig. 16 is the single line diagram of the IEEE-24 bus one-area test system, adapted from [32]. It is composed of 24 buses, 34 power lines, 10 generation stations, and 17 load points. Its total generating capacity is 3405MW and a varying load which annual peak is 2850MW. The total generating capacity and load are distributed across the network as described in [33]. The buses are assumed perfectly reliable and the transmission lines, binary-state. We retain the failure and repair characteristics of the transmission lines but modify a few other properties to make the system more realistic and compatible with the proposed approach. These modifications are summarised thus;

- Multiple generation units at a bus, have been represented by a single unit with a generating capacity equivalent to the sum of the generating capacities of the units.
- To make the network more sensitive to the unavailability of transmission lines and generation units, the maximum transmission capacities of the former and minimum allowable loads of the latter are considered in the analysis. These capacities and limits, are respectively given in [33] and [32]. Please note that the minimum load for the unit at bus 22 is set to 25MW instead of the 300MW suggested in [32]. The reason for this is, its contribution

TABLE VII
MAINTENANCE DATA FOR GENERATION UNITS

Gen. Type	Bus Number	Spare Usage Prob.		PM		Transition Distribution Parameters				
		CM	PM	Interval	Duration	1-2	2-1	2-3	1-3	3-1
1	22	0.7	0.9	1200	U(156,180)				Wb(2234,2)	Exp(20)
2	1 & 2	0.9,0.25	0.9	1200	U(60,66)	Exp(980)	Exp(20)	Wb(1106,2.3)	Wb(2212,2)	Exp(40)
3	7	0.8,0.4	0.9	1200	U(60,66)	Exp(600)	Exp(25)	Wb(677,2.3)	Wb(1354,2)	Exp(50)
4	15,16 & 23	0.8,0.3	0.9	1000	U(81,87)	Exp(480)	Exp(20)	Wb(542,2.3)	Wb(1083,2)	Exp(40)
5	13	1.0,0.5	0.9	1000	U(102,108)	Exp(575)	Exp(50)	Wb(649,2.3)	Wb(1298,2)	Exp(100)
6	18 & 21	1,0.6	0.9	1000	U(123,129)	Exp(550)	Exp(75)	Wb(621,2.3)	Wb(1241,2)	Exp(150)

TABLE VIII
MAINTENANCE COSTS FOR GENERATION UNITS

Gen. Type	CM		PM	
	CS	CPHM	CS	CPHM
1	180	20	108	12
2	180	20	108	12
3	180	20	108	12
4	200	25	120	15
5	280	40	168	24
6	300	50	180	30

to the total load when every component works correctly is only about 37.5MW. A minimum allowable load of 300MW, therefore, would mean it operates only on failure of another unit. This, in other words, reduces the unit to cold standby, thereby defeating our intention of making every component useful to the system throughout the mission.

- The buses are assigned maximum capacities according to the following rules;

  1) For load and generation buses, the maximum capacity is arbitrarily set to 3 times the capacity of the generation unit or load.
  2) For buses with both a generation unit and load, the capacity is set to 3 times the generating capacity or load, whichever is greater.
  3) For all other buses, the capacity is set to 3 times the maximum of the capacities of the buses they are connected to.

- Each generation unit, with the exception of the unit at bus 22, is assumed to exist at three possible distinct output levels; 100%, 50%, and 0% of its rated capacity. Unit 22 operates at only two levels; 100% and 0% rated capacity.

1) Maintenance Information: The failure times of the transmission lines are exponentially distributed. As a consequence, they undergo corrective maintenance only, with an assumed 0.9 likelihood of spares being used. Due to their less bulkiness, it's assumed the maintenance crew are able to carry with them these spares. The maintenance of the lines, therefore, is immune to delays in the availability of spares.

The generation units, on the other hand, undergo both PM and CM, and are susceptible to all the operational dynamics described in Section II. Table VII contains their failure and maintenance parameters, where states 1, 2, and 3 respectively represent nominal performance, partial, and complete failure. Their replacement probability during CM is represented by a pair which elements respectively define the probabilities

associated with states 3 and 2. Where applicable, the diagnosis and CM durations have the same distribution, with means in the ratio, $1 : 4$. For instance, the transition of the unit at bus 13 from state 3 to 1, denoting repairs from complete failure, is exponentially distributed with mean 100. Therefore, the diagnosis and CM durations are also exponentially distributed with means 20 and 80 respectively. All transition times are in hours and k_i for generation units is conservatively assumed to be 0.3. Please also note that the data presented in Table VII are for illustrative purposes only.

2) Maintenance Grouping and Costs: The network components are arranged into three maintenance groups, and each group maintained by a separate maintenance company. The transmission lines above buses 11, 12, and 24 make maintenance group 1, the remaining lines, group 2, and the generation units constitute group 3. Each maintenance team in groups 1 and 2 is paid a fixed £5 per hour and a fixed £100 per successful maintenance action. Teams in group 3 earn £8 every hour and £120 for every successful maintenance action. Due to economic constraints, the operator imposes the total number of maintenance teams to not exceed 16. The cost of one transmission line spare is averaged at £150, the cost per hour of transmission line maintenance, at £15, and the cost levels for the generation units, as defined in Table VIII.

3) Objective: The current maintenance strategy, hereafter referred to as the base strategy, assumes CM of partially failed components and PM can be initiated at any time, subject to the availability of maintenance teams. For one annual load cycle of 8736 hours (see [33]) and £100 per MWh of electricity consumed, we determine the optimal maintenance team size for this strategy and compare its effectiveness with three complex strategies. The base strategy, for simplicity, is labelled strategy 1, and the complex strategies, as outlined thus;

- Strategy 2: PM, and CM of partially failed generation units only when they are not required.
- Strategy 3: PM, and CM of partially failed generation units only when system performance is nominal.
- Strategy 4: PM of generation units only when system performance is nominal but CM of partially failed units can be carried out at any time.

Each maintenance strategy is computed for the case when the units,

- (a) Are kept out of operation during maintenance suspensions
- (b) Are returned into operation during maintenance suspensions

4) System modelling: Since the goal is to identify the optimal maintenance strategy, a DC flow analysis, using the procedure proposed in [22], is employed to compute the system reliability and performance indices. The buses, generation units, and load points are modeled as nodes, while the transmission lines are modeled as edges, in the system graph model. In this case-study, we have retained the edge attribute of the transmission lines, to keep the number of nodes moderate and improve performance. Consequently, the vector of maximum edge capacities is modified after every transition involving a transmission line, and both this vector and the vector of node capacities are required for system flow calculation. Fig. 17 (a) shows the graph model of the system, where Un and Ln respectively denote the generation unit and load point at bus n. Fig. 17 (b) is the same graph but with only one edge of each reciprocal pair [22] shown for clarity. In both cases, the number along each edge defines the maximum flow along that edge, as a fraction of the annual peak load.

The effective EENS of the system (given the multiple load points) could be computed as proposed in Section II. However, the computation is rendered less complicated by representing the global system output by a virtual node which flow is the sum of the flows through all 17 load points. The flow history of this virtual node is recorded during simulation, and subsequently used to compute the effective EENS, instead of all 17 nodes. Being mindful of the computational demand of simulation algorithms, we employ a smart procedure to treat the variable demand on the system. Recall the objective of system reliability analysis is to determine the maximum achievable system performance as a consequence of component failure and maintenance. For this reason, we obtain the instantaneous system performance, $Y(t)$, assuming the demand is fixed at its peak annual value. However, under this assumption, the system is no longer strictly demand-driven (since the actual demand varies with time), and $Y(t)$ has to be normalized to make it compatible with (1) and (2). The normalization entails expressing $Y(t)$ as a function of the same time-step as the instantaneous demand, $Y_d(t)$, such that they both have equal lengths, and applying the following;

$$Y(t) = min\{Y(t), Y_d(t)\} \qquad (17)$$

Normally, variable demand is treated by performing the simulation with respect to the time-step defined by the demand and the events generated by component failures and maintenance. It is, therefore, easy to deduce the computational efficiency of the procedure employed here, relative to the widely practised. The procedure is correct for all single load point systems, as well as multiple load-point systems where the quantity of interest is the total output, and not the output through the individual load points.

To derive the set, $\mathbb{N}$, of possible maintenance team combinations, we ignore the possibility of a 0 maintenance team in any of the maintenance groups. This is due to the fact that we already know (from the previous case-study) non-repairable maintenance strategies to be grossly inefficient. Recall also that maintenance groups 1 and 2 are composed of equal number of components with the same failure and repair characteristics. They, therefore, have the same optimal

TABLE IX
OPTIMAL SYSTEM LOSS AS A FUNCTION OF MAINTENANCE STRATEGY

Strategy		EENS(%)	L($£10^6$)	Optimal number of teams		
				Group 1	Group 2	Group 3
1	a	0.3940	6.6324	1	1	3
	b	0.2468	4.4712	2	2	3
2	a	2.4218	37.6617	1	1	4
	b	2.4780	38.6764	3	3	4
3	a	1.3592	21.3563	1	1	3
	b	1.5049	23.6498	1	1	4
4	a	0.3373	5.9026	1	1	5
	b	0.2128	3.9513	2	2	3

maintenance team size. Given these constraints and the upper bound imposed by the operator on the total number of maintenance teams, $\mathbb{N}$ contains 50 maintenance team combinations.

5) Component Modelling: Figs. 18 and 19 are the system's simplified component models, showing only the required transitions, as discussed in Section III-E. Since the transmission lines are not susceptible to maintenance interruptions, their failure diagnosis and actual repair have been collectively represented by the CM state. This, however, implies the number of spares used cannot be directly obtained from the simulation, as spares used are accounted for only if the component enters Diagnosis or PM state (see Algorithm 2). The total spares used, therefore, is obtained from the product of the spares usage probability and the number of CM to W transitions. Please note the models in Figs. 18 and 19 are based on the assumption that components are kept out of operation during maintenance suspensions. Those for the case when components are returned into operation can be easily deduced from Figs. 2 to 4. It is also worthwhile noting that, the simplified component models for regimes 1 to 3 of Section II-C are equivalent.

6) Results and Discussions: The system was analysed on the same computer used for the previous case-study, and the outcome is summarized in Table IX. The table provides the EENS as a percentage of the total expected output, the expected loss, and the optimal maintenance team combination, for each strategy. Each sample of a candidate solution took an average of 0.8 seconds, using 10 MATLAB workers. Given the large number of candidate solutions, the number of samples per candidate solution was set to 500. The sensitivity of the optimal solution to the costs considered in the previous case-study and a few other costs, was also investigated. The additional costs considered are;

- Cost per hour of CM and cost per CM call (CPHM1).
- Cost per hour of PM and cost per PM call (CPHM2).
- Total maintenance cost (MC); a combination of FMC, CPHM1, CPHM2, and the cost per CM and PM call.
- All costs relevant to the system loss function (ALL).

Deducing from the data in Table IX, the optimal maintenance strategy is strategy 4 (b). In this strategy, CM of partially failed generation units can be initiated at any time, but PM, only when system performance is nominal, with components returned into operation during maintenance suspensions (see the beginning of this subsection). Postponing both CM and PM until component shut-down (strategy 2) appears to be the most inefficient, contrary to what obtained in the previous
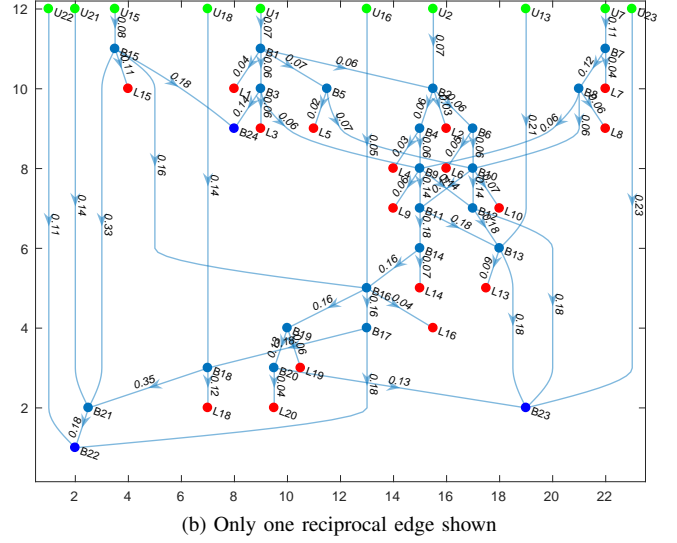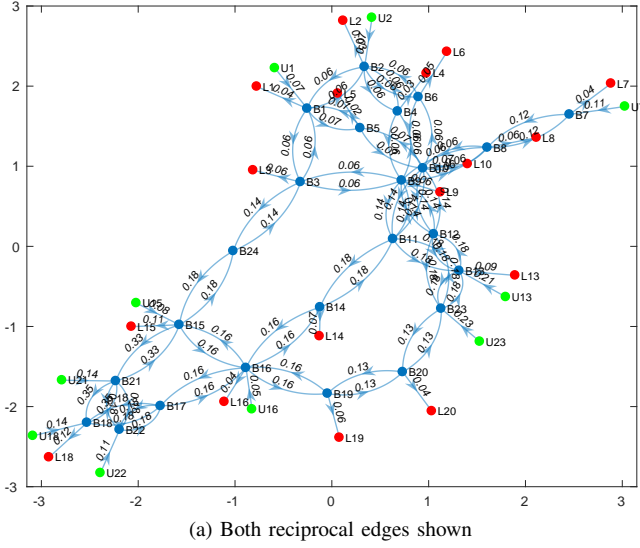
(a) Both reciprocal edges shown

(b) Only one reciprocal edge shown

Fig. 17. System graph model

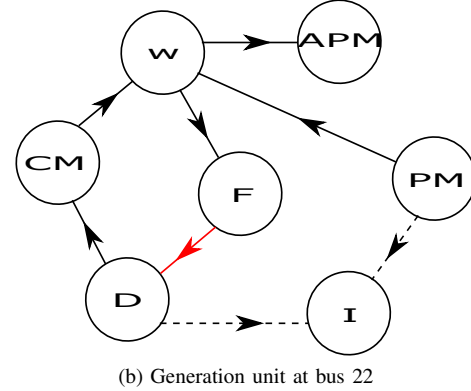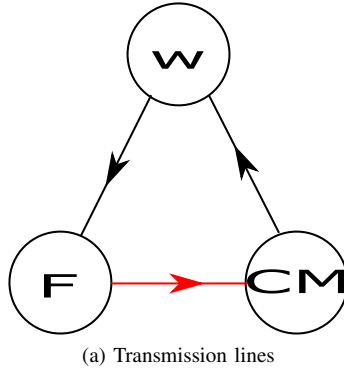(a) Transmission lines

(b) Generation unit at bus 22

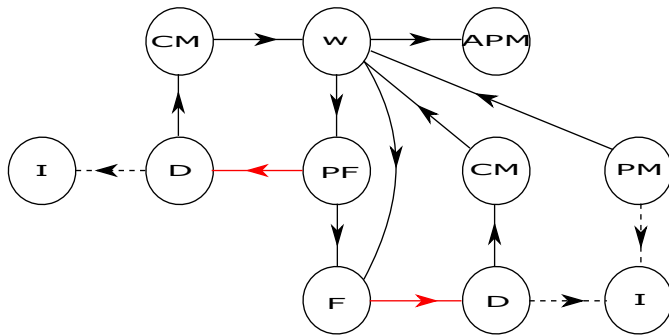Fig. 18. Simplified multi-state model for binary-state components

Fig. 19. Simplified multi-state model for multi-state generation units

case-study. This observation reiterates the point that the optimality of a given maintenance strategy depends on specific properties of the system. For $0 \leq k_f \leq 100$, strategy 4 (b) remains optimal, but the optimal number of maintenance teams varies as depicted by Fig. 20. It should be noted that cost parameters with no effect on the optimal number of maintenance teams have been left out in Figs. 20 (a) and 20

(b). Given maintenance groups 1 and 2 are made up of the transmission lines only (which do not undergo PM), CPHM and CPHM1 are equivalent, explaining the absence of CPHM1 and CPHM2 in Fig. 20 (a). A notable conclusion drawn from Fig 20 is that the optimal number of maintenance teams is most affected by the cost of electricity (EC) and the fixed cost per maintenance team (FMC). It is also easily deducible that the number of teams required for optimality reduces and increases with reduction in EC and FMC respectively, both observations conforming to common reasoning.

Fig. 21 shows the variation in system loss with changes in cost levels in the range, $0 \leq k_f \leq 2$. For clarity, system response over the ranges $0 \leq k_f \leq 1$ and $1 \leq k_f \leq 2$ has been presented separately in Figs. 21 (a) and 21 (b) respectively. With $k_f = 1$ as reference, Fig. 21 (a) defines the sensitivity of the total system loss to cost reductions and Fig. 21 (b), to cost increments. In both cases, the cost of electricity and the overall maintenance cost impact system loss the most. However, the system shows very little sensitivity to both the cost of spares and the cost per hour of PM action, suggesting a few PM actions and low spares usage. The low system loss sensitivity
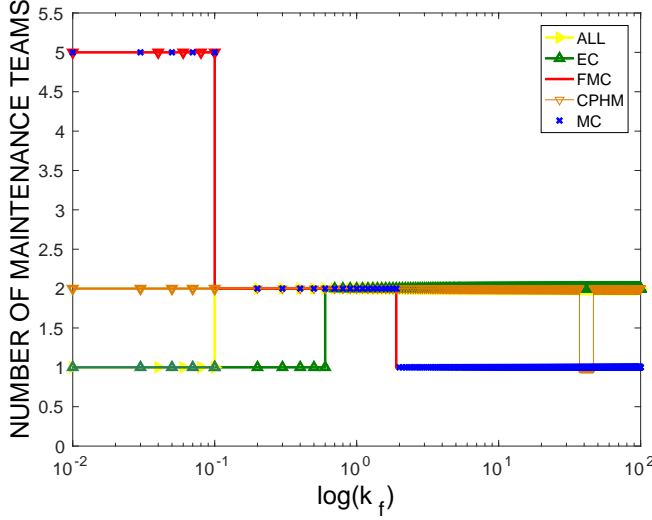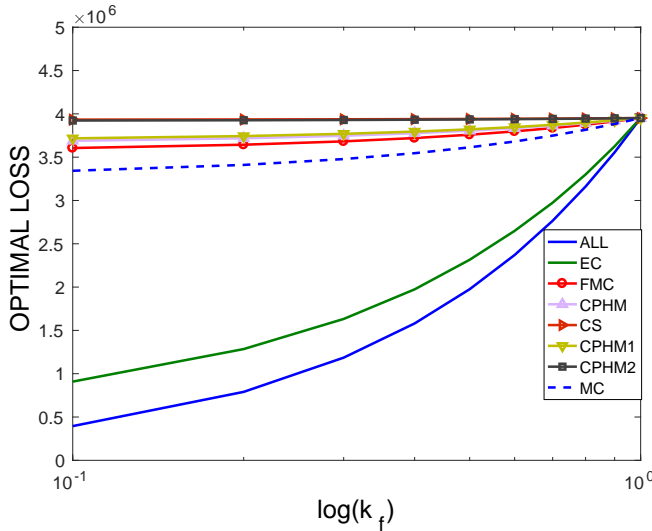
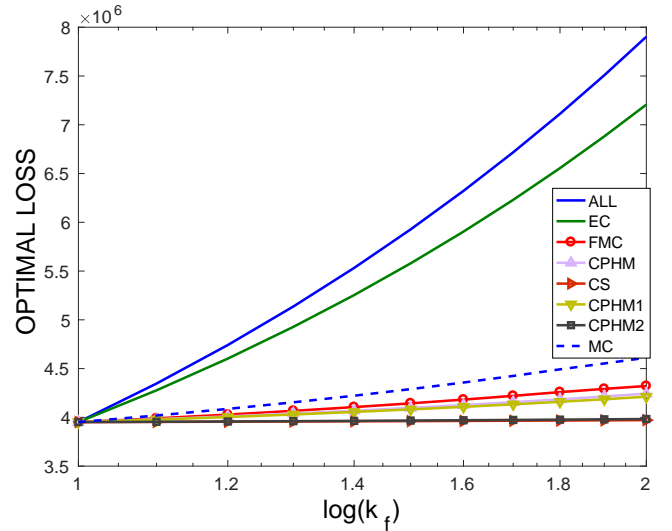(a) Groups 1 and 2
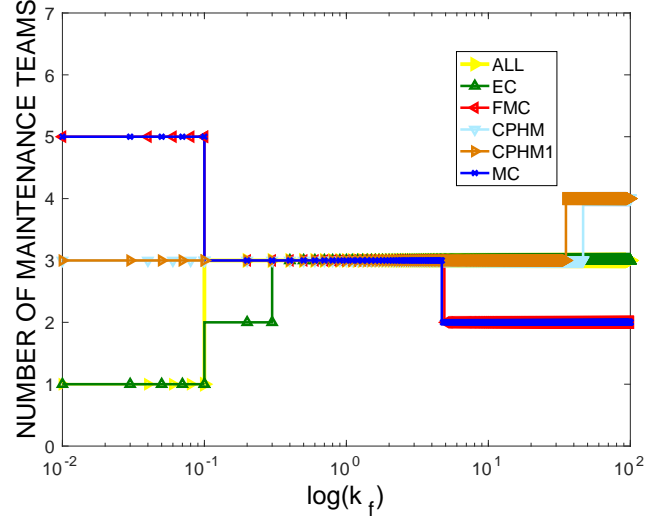
(b) Group 3

Fig. 20.   Optimal maintenance team sensitivity to cost levels

(a) Cost reduction

(b) Cost increment

Fig. 21.   System loss sensitivity to cost levels

to CPHM2 is explained by the fact that only 10 of the 44 system components undergo PM. Given strategy 4 imposes PM be initiated only if system performance is nominal, a good number of these components fail before their PM commences.

V. CONCLUSIONS

It is realistic to think increasing the number of maintenance teams improves the performance and reliability of a multi-component system. However, a threshold exists exceeding which no gains are realised. Rather, it results in increased operational costs, borne from the imbalance between income and expenditure. This threshold, as expected, varies with the maintenance strategy, the input costs to the system's cost

model, the topology of the system, and the non-topological functional relationships between its components.

In this work, a maintenance strategy optimization framework, aiding proper maintenance scheduling and robust maintenance decisions has been presented. Applicable to both binary and multi-state systems of any structure, the framework proposes a multi-state model to define the behaviour of components under various maintenance strategies. A non-system-specific, event-driven Monte Carlo simulation, based on the load-flow approach proposed in [22] is employed to replicate the operation of the system. This simulation algorithm, together with the multi-state component model, enhances the implementation of complex maintenance strategies. For instance, a component may belong to two maintenance groups practising dedicated and shared maintenance respectively. It

is also possible to have the situation where the preventive maintenance of a certain group of components can be initiated as soon as they are due, and another group, delayed to coincide with a shut down event or nominal system performance. Many more contrasting combinations of regimes are possible, without the need to modify the simulation algorithm. The framework is also built on a cost model structured to allow the sensitivity analysis of the optimal solution from a single reliability evaluation. These attributes render it novel, efficient, and generally applicable to power and other systems alike.

The framework has been successfully used to optimize the maintenance strategies for two realistic power systems, obtaining insightful information on their maintenance. The relationship derived between the optimal number of maintenance teams and the cost of electricity, for instance, is a very useful tool, given a volatile electricity market. The framework, therefore, can shape the quality of maintenance-related decisions, even in the presence of external dynamics.

ACKNOWLEDGMENT

REFERENCES

[1] E. Reihani, A. Sarikhani, M. Davodi, and M. Davodi, "Reliability based generator maintenance scheduling using hybrid evolutionary approach," *International Journal of Electrical Power & Energy Systems*, vol. 42, no. 1, pp. 434 – 439, 2012.

[2] R. Dekker, "Applications of maintenance optimization models: a review and analysis," *Reliability Engineering & System Safety*, vol. 51, no. 3, pp. 229 – 240, 1996, maintenance and reliability.

[3] M. Marseguerra and E. Zio, "Optimizing maintenance and repair policies via a combination of genetic algorithms and monte carlo simulation," *Reliability Engineering & System Safety*, vol. 68, no. 1, pp. 69 – 83, 2000.

[4] E. Borgonovo, M. Marseguerra, and E. Zio, "A monte carlo methodological approach to plant availability modeling with maintenance, aging and obsolescence," *Reliability Engineering & System Safety*, vol. 67, no. 1, pp. 61 – 73, 2000.

[5] J. Heo, M. Kim, and J. Lyu, "Implementation of reliability-centered maintenance for transmission components using particle swarm optimization," *International Journal of Electrical Power & Energy Systems*, vol. 55, pp. 238 – 245, 2014.

[6] A. C. Marquez and A. S. Heguedas, "Models for maintenance optimization: a study for repairable systems and finite time periods," *Reliability Engineering & System Safety*, vol. 75, no. 3, pp. 367 – 377, 2002.

[7] H. C. Vu, P. Do, A. Barros, and C. Berenguer, "Maintenance grouping strategy for multi-component systems with dynamic contexts," *Reliability Engineering & System Safety*, vol. 132, pp. 233 – 249, 2014.

[8] B. Yssaad and A. Abene, "Rational reliability centered maintenance optimization for power distribution systems," *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 350 – 360, 2015.

[9] G. Levitin and A. Lisnianski, "Optimization of imperfect preventive maintenance for multi-state systems," *Reliability Engineering & System Safety*, vol. 67, no. 2, pp. 193 – 203, 2000.

[10] W. Li and H. Pham, "An inspection-maintenance model for systems with multiple competing processes," *Reliability, IEEE Transactions on*, vol. 54, no. 2, pp. 318–327, June 2005.

[11] Y. Liu and H.-Z. Huang, "Optimal replacement policy for multi-state system under imperfect maintenance," *Reliability, IEEE Transactions on*, vol. 59, no. 3, pp. 483–495, Sept 2010.

[12] C. M. Tan and N. Raghavan, "A framework to practical predictive maintenance modeling for multi-state systems," *Reliability Engineering & System Safety*, vol. 93, no. 8, pp. 1138 – 1150, 2008.

[13] V. Makis and A. K. Jardine, "A note on optimal replacement policy under general repair," *European Journal of Operational Research*, vol. 69, no. 1, pp. 75 – 82, 1993. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0377221793900922

[14] W. Hongzhou and P. Hoang, "Maintenance policies and analysis," in *Reliability and Optimal Maintenance*, ser. Springer Series in Reliability Engineering. Springer London, 2006, pp. 31–49.

[15] J. J. McCall, "Maintenance policies for stochastically failing equipment: A survey," *Management Science*, vol. 11, no. 5, pp. 493–524, 1965. [Online]. Available: http://www.jstor.org/stable/2627585

[16] R. Barlow and F. Proschan, *Mathematical Theory of Reliability*. New York: Wiley, 1965.

[17] A. Lisnianski, I. Frenkel, L. Khvatskin, and Y. Ding, "Maintenance contract assessment for aging systems," *Quality and Reliability Engineering International*, vol. 24, no. 5, pp. 519–531, 2008. [Online]. Available: http://dx.doi.org/10.1002/qre.913

[18] J. Barata, C. Soares, M. Marseguerra, and E. Zio, "Simulation modelling of repairable multi-component deteriorating systems for on-condition maintenance optimisation," *Reliability Engineering & System Safety*, vol. 76, no. 3, pp. 255 – 264, 2002.

[19] M. Marseguerra, E. Zio, and L. Podofillini, "Condition-based maintenance optimization by means of genetic algorithms and monte carlo simulation," *Reliability Engineering & System Safety*, vol. 77, no. 2, pp. 151 – 165, 2002.

[20] M. Malhotra and K. S. Trivedi, "Dependability modeling using petri-nets," *IEEE Transactions on Reliability*, vol. 44, no. 3, pp. 428 – 440, 1995. [Online]. Available: http://dx.doi.org/10.1109/24.406578

[21] S. Dhople, L. DeVille, and A. Dominguez-Garcia, "A stochastic hybrid systems framework for analysis of markov reward models," *Reliability Engineering & System Safety*, vol. 123, pp. 158 – 170, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0951832013002950

[22] H. George-Williams and E. Patelli, "A hybrid load flow and event driven simulation approach to multi-state system reliability evaluation," *Reliability Engineering & System Safety*, vol. 152, pp. 351 – 367, 2016.

[23] E. Zio, P. Baraldi, and E. Patelli, "Assessment of the availability of an offshore installation by monte carlo simulation," *International Journal of Pressure Vessels and Piping*, vol. 83, no. 4, pp. 312 – 320, 2006. [Online]. Available: http://dx.doi.org/10.1016/j.ijpvp.2006.02.010

[24] H. George-Williams and E. Patelli, "Efficient availability assessment of reconfigurable multi-state systems with interdependencies," *Reliability Engineering and System Safety*, vol. 15, pp. 431–444, 2017.

[25] B. Tuffin, P. Choudhary, C. Hirel, and K. Trivedi, "Simulation versus analytic-numeric methods: a petri net example," in *Proc. of the 2nd VALUETOOLS Conference*, 2007.

[26] M. Bouissou, H. Elmqvist, M. Otter, and A. Benveniste, "Efficient monte carlo simulation of stochastic hybrid systems," in *Proceedings of the 10th International Modelica Conference; March 10-12; 2014; Lund; Sweden*, no. 96. Linkoping University Electronic Press; Linkopings universitet, 2014, pp. 715–725.

[27] A. B. Huseby and B. Natvig, "Discrete event simulation methods applied to advanced importance measures of repairable components in multistate network flow systems," *Reliability Engineering & System Safety*, vol. 119, pp. 186 – 198, 2013.

[28] J.-A. Li, Y. Wu, K. K. Lai, and K. Liu, "Reliability estimation and prediction of multi-state components and coherent systems," *Reliability Engineering & System Safety*, vol. 88, no. 1, pp. 93 – 98, 2005.

[29] M. Kojima, S. Mizuno, and A. Yoshise, "A primal-dual interior point algorithm for linear programming," in *Progress in Mathematical Programming*, N. Megiddo, Ed. Springer New York, 1989, pp. 29–47.

[30] E. Patelli, *Handbook of Uncertainty Quantification*. Springer International Publishing, 2017, ch. COSSAN: A Multidisciplinary Software Suite for Uncertainty Quantification and Risk Management, pp. 1–69.

[31] E. Patelli, M. Broggi, M. D. Angelis, and M. Beer, "Opencossan: An efficient open tool for dealing with epistemic and aleatory uncertainties," in *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management - Proceedings of the 2nd International Conference on Vulnerability and Risk Analysis and Management, ICVRAM 2014 and the 6th International Symposium on Uncertainty Modeling and Analysis, ISUMA 2014*, 2014, pp. 2564 – 2573. [Online]. Available: http://dx.doi.org/10.1061/9780784413609.258

[32] C. Ordoudis, P. Pinsona, J. M. Morales, and M. Zugno, "An updated version of the ieee rts 24-bus system for electricity market and power system operation studies," Technical University of Denmark, Tech. Rep., 2016.

[33] Probability Methods Subcommittee, "IEEE reliability test system," *IEEE Transactions on power apparatus and systems*, no. 6, pp. 2047–2054, 1979.

Hindolo George-Williams obtained a BEng(Hons) degree in Electrical/Electronic Engineering from the University of Sierra Leone in 2010 and an MSC(Eng) in Energy Generation from the University of Liverpool, three years later. He is currently enrolled on the dual PhD degree program between the University of Liverpool and the National Tsing Hua University in Taiwan. His PhD research is focused on the Probabilistic Risk Assessment of nuclear power plants. In 2012, Hindolo received the best project award from the Sierra Leone Institute of Engineers in recognition of his outstanding execution of his final BEng project. He also worked as Maintenance Engineer (for a period of 30 months) for the Sierra Leone affiliate of the French oil giant, TOTAL.

Edoardo Patelli is a member of the Institute for Risk and Uncertainty at the University of Liverpool, UK and a honorary member of the National Tsing Hua University, Taiwan. He is a Nuclear Engineering graduate from the Politecnico di Milano (Italy) and carried out his doctoral work in Radiation Science and Technology at the same Institute in the group of Professor Marzio Marseguerra and Enrico Zio. He then moved as a research associate to the University of Innsbruck (Austria) in the group of Professor Schuëller. Dr Patelli is co-Principal investigator of the Centre for Doctoral Training in Quantification and Management of Risk & Uncertainty in Complex Systems & Environments and a member of the Centre for Doctoral Training in "Next-Generation-Nuclear". Edoardo has published more than 120 contributions in international journals and proceedings of international conferences. He has supervised more than 20 PhD students on site and in collaboration with international partners. He is a guest-editor of international journals (e.g., the International Journal of Reliability & Safety and Structural Safety) and editorship of Springer's Encyclopaedia of Earthquake Engineering. He has also organised multi-disciplinary international conferences on risk and vulnerability (e.g., ASCE-ICVRAM-ISUMA 2014, IPW2015) and a number of Mini-symposia in different international conferences.