Data-Driven Maintenance and Operations Scheduling in Power Systems under Decision-Dependent Uncertainty

Beste Basciftci, Shabbir Ahmed, Nagi Gebraeel

H. Milton Stewart School of Industrial and Systems Engineering

Georgia Institute of Technology

Abstract

Generator maintenance scheduling plays a pivotal role in ensuring uncompromising operations of power systems. There exists a tight coupling between the condition of the generators and corresponding operational schedules, significantly affecting reliability of the system. In this study, we effectively model and solve an integrated condition-based maintenance and operations scheduling problem for a fleet of generators with an explicit consideration of decision-dependent generator conditions. We propose a sensor-driven degradation framework with remaining lifetime estimation procedures under time varying load levels. We present estimation methods by adapting our model to the underlying signal variability. Then, we develop a stochastic optimization model that considers the effect of the operational decisions on the generators' degradation levels along with the uncertainty of the unexpected failures. As the resulting problem includes nonlinearities, we adopt piecewise linearization along with other linearization techniques and propose formulation enhancements to obtain a stochastic mixed-integer linear programming formulation. We develop a decision-dependent simulation framework for assessing the performance of a given solution. Finally, we present computational experiments demonstrating significant cost savings and reductions in failures in addition to highlighting computational benefits of the proposed approach.

Keywords: Condition monitoring, maintenance, stochastic optimization, unit commitment, power systems.

1 Introduction

Ensuring reliable and cost effective operations of generators is an important problem in power systems. One of the key factors that impacts this problem is related to generator maintenance scheduling, since resulting schedules determine availability and reliability of the generators. Most classical maintenance approaches rely on predetermined time intervals or safety margins for scheduling maintenance. Consequently, these methods often result in excessive or unnecessary maintenance events, especially when designed in a conservative manner. They also do not provide much needed visibility into the actual condition of the generators, and thus, still experience unexpected failures. The emergence of sensor technology has enabled the incorporation of generator state-of-health into maintenance and operations scheduling. Operational decisions, such as on-off actions for generators and their dispatch amount, play a pivotal role in this regard as higher (lower) loads result in an accelerated (decelerated) degradation process, which may require scheduling maintenance at an earlier (or later) time. Thus, the effect of operational decisions on the corresponding component's aging is critical in determining its availability. To address these issues, we propose an optimization framework for condition-based generator maintenance and operations scheduling that accounts for the loading profiles derived from the operational decisions.

Generator maintenance scheduling constitutes an important class of problems in power systems, see Froger et al. (2016) for a recent survey. The literature on maintenance scheduling can be categorized into three groups. The first group of studies focuses on periodic maintenance routines that are based on a predetermined time schedule (Billinton and Mo (2005); Conejo et al. (2005)) with additional constraints that inform maintenance decisions, such as enforcing one maintenance per year. These studies neglect potential information related to the condition/performance of the generators. The second line of research (Abiri-Jahromi et al., 2012; Pourahmadi et al., 2017; Levitin et al., 2017) adopts a reliability-based approach by considering failure rates and reliability metrics, such as mean-time-to-failure for scheduling maintenance. These approaches tend to adopt a general schedule across all generators regardless of their unit-to-unit variations, whether in the way they are operated or the manner in which they degrade. A third group of recent studies (Yildirim et al., 2016b, 2016a; Basciftci et al., 2018) considers generator degradation by monitoring cumulative damage and other forms of wear and tear using sensor technology. These studies focus on leveraging real-time generator-specific degradation signals to estimate statistical distributions of the generator's remaining lifetime. The predictions can be updated in real-time and leveraged when solving the operational problem.

However, these models stop short of modeling the impact of operational decisions on generator degradation (hereafter referred to as *load-independent* models). In reality, generator loading has a significant effect on how fast a generator degrades. Harsh operating and loading conditions, i.e., dispatch decisions, tend to accelerate physical degradation. Thus, there is a tight coupling between the operational decisions and its effect on generator degradation.

Data-driven degradation modeling has played a key role in predicting remaining lifetime of machines and capital-intensive assets. A review of various modeling approaches used for this purpose can be found in Si et al. (2011). The relationship between loading conditions and equipment reliability has been a well-studied area in the reliability literature. Traditionally, operating conditions have been considered as model covariates in classic location-scale models and proportional hazard models, see Kumar and Klefsjo (1994). The accelerated degradation testing framework proposed by Doksum and Hóyland (1992) is among the early models that considers the interaction between the stress conditions and degradation rates. The authors model degradation using a Brownian motion and present a time-scale transformation proportional to the testing stress. Recently, time-varying operating conditions are integrated with sensor-based degradation models for predicting remaining life distributions in an adaptive manner using a Bayesian framework, see Gebraeel and Pan (2008); Bian and Gebraeel (2013); Liao and Tian (2013).

Although load-dependent degradation modeling has been studied in the reliability literature, integrating it into maintenance optimization has not been explored in depth. Ideally, effective maintenance scheduling should not only consider the health of an asset, but also its loading/operating conditions. In power systems, operating conditions of the generators are determined by the unit commitment (UC) problem, which is a well-studied problem in literature (see Zheng et al. (2015); van Ackooij et al. (2018) for recent reviews). The UC problem can be also referred as the operational problem in power systems. This problem aims to determine which generators will be on or off and how much energy they need to produce by considering restrictions of the underlying power network such as generation capacities, transmission line limits, etc. As the loading/operating conditions are critical in assessing generators' conditions, the operational problem and corresponding load-dependent degradation modeling need to be incorporated into maintenance scheduling. With the exception of recent work by Yildirim et al. (2019), there has not been a formal framework for using load-dependent degradation models in scheduling maintenance. In Yildirim et al. (2019), the authors solve the UC problem by considering the effects of generator loads on degradation. They categorize generator loading into three levels and jointly optimize operations and maintenance decisions. Although the authors consider the stochasticity of the degradation process, they neither formally account for unexpected failure possibilities nor the continuous load amounts within the optimization model.

Most of the power system problems involve uncertainties which can only be addressed by a stochastic optimization framework. These uncertainties also play a critical role in the joint scheduling of maintenance and operations. Several studies (Canto, 2008; Wu et al., 2008, 2010) consider uncertainty of price and demand for this scheduling problem. Most recently, Basciftci et al. (2018) proposes a stochastic optimization framework for modeling sensor-based generator maintenance and operations scheduling problem by explicitly considering generator failures through scenarios and chance constraints. Nevertheless, the proposed approach does not consider the effect of the operational decisions on generators' degradation. Neglecting this effect results in a simplified prognostics procedure and a mixed-integer linear optimization model to determine maintenance and operations schedules. Modeling the dependency between operational decisions and degradation processes involve decision-dependent (endogenous) uncertainties. In decision-dependent uncertainty, the distribution of a random variable changes due to the planner's decisions, and complex modeling procedures are required for its correct representation (see Goel and Grossmann (2006)). In our case, the dispatch levels of generators affect the distribution of the generators' remaining lifetimes, resulting in an endogenous uncertainty to represent unexpected failure possibilities, which need to be considered when scheduling maintenance and operations.

In this paper, we develop a novel stochastic optimization framework that determines the maintenance and operations schedules of the generators while considering the impact of the decision-based degradation amounts. We refer to our comprehensive approach as *load-dependent* in the rest of this paper. Our main contributions can be summarized as follows:

1. We formulate a stochastic optimization model that jointly optimizes maintenance and operations while accounting for the endogenous effect of operational decisions on degradation. The resulting model includes nonlinearities due to the decision-dependent structure of the cumulative distribution functions of the remaining lifetimes and maintenance cost functions associated with generators. We develop linearization procedures for these nonlinearities to obtain a stochastic mixed-integer linear programming formulation, and propose formulation enhancements. We also extend the chance constraint proposed in Basciftci et al. (2018) to the decision-dependent setting.

2. We build on existing literature and propose a data-driven degradation modeling framework

for capturing the effect of the continuous loading profiles (rather than discrete levels) that are functions of the generators' minimum and maximum production capacities. We also develop an estimation method to consider the effect of the load decisions by taking into account the signal variability.

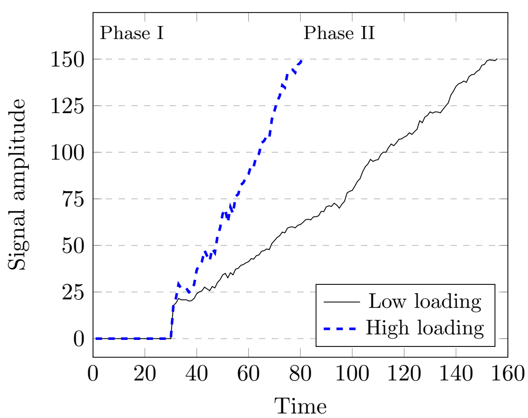
3. We propose a decision-dependent simulation framework to evaluate the performances of resulting maintenance and operations schedules. We provide a comprehensive computational study on representative IEEE test cases by comparing the proposed load-dependent approach with load-independent approaches under different congestion levels and conservativeness amount of the chance constraint. Our experiments demonstrate the success of load-dependent schedules resulting in significant cost savings of up to 20% and failure preventions. Finally, we provide experiments demonstrating the computational efficiency of the formulation improvements compared to generic methods with more than 17 times speedup.

The remainder of the paper is organized as follows: Section 2 discusses the degradation modeling framework along with the data-driven estimation procedures. Section 3 presents the optimization methodology for modeling load-dependent condition-based maintenance and operations schedules with our proposed enhancements. Section 4 illustrates the computational results by developing a simulation procedure, and the efficiency of the proposed approach from various aspects. Section 5 concludes the paper with final remarks.

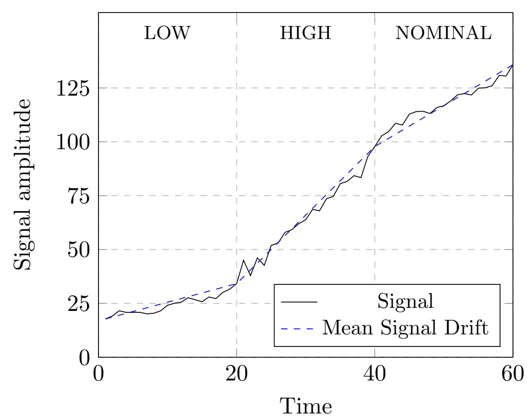
2 Degradation Modeling

Unexpected failures of power generators can have catastrophic consequences. Thus, monitoring the state-of-health of power generators is necessary to maintain their availability and improve their reliability. Condition monitoring, the process of using sensors to assess the state-of-health of a machine, is becoming more prevalent across numerous industrial applications. Raw sensor data from power generators can be synthesized into degradation signals that represent the severity of the underlying physical degradation processes taking place in the generator. The degradation state of the generators and their remaining lifetime distributions can be estimated using degradation signals.

One of the most critical factors that determines the degradation of a generator is severity of its operating condition. Generators tend to degrade faster when operated close to their maximum capacity. To demonstrate this effect, we illustrate degradation signals in Figure 1 that follow a Brownian motion with drift to mimic the degradation of a rotating machinery under varying loading conditions. Figure 1a shows an example of two degradation signals under low and high loading. Phase I is considered an "as good as new" state with no signs of degradation whereas Phase II highlights the progressive nature of physical degradation and its manifestation in a gradually increasing degradation signal. Figure 1b represents potentially how the degradation rate changes with low and high loading conditions. The dotted line in Figure 1b shows the mean drift of the degradation signal under each load condition. The plot illustrates how the degradation rate at low loading from time 0 to 20 is lower than the nominal condition. The converse is true for the high load from time 20 to 40. This example highlights the importance of capturing the effects of operational decisions on generator conditions when solving the maintenance and operations scheduling problem.



(a) Effect of load decisions on two signals.



(b) Effect of time varying load levels.

Figure 1: Illustration of the effect of load decisions on degradation signals.

Usually, an equipment operates for a period of time without any signs of degradation (Phase I). This phase is often random and hypothetically equipment does not fail in this period due to degradation. However, equipment is often subject to external factors such as human related operational errors that can cause an unexpected failure event. These failures are rare and not related to equipment's degradation, and often cannot be predicted. Phase II is characterized by a degradation process that is significant enough to be observed using some form of sensor technology. Most machines and equipment can operate in these partially degraded modes for a significant period of time. In fact, Phase II is typically where degradation trends can be leveraged when predicting remaining lifetime. Our analysis focuses on the second phase of a degradation process for modeling failures due to degradation.

2.1 Load-dependent degradation model

We model the degradation signal of a generator i as a continuous-time continuous-state model, denoted by $\{S_i(t) : t \geq 0\}$. We assume that $S_i(t)$ has the following functional form;

$$S_i(t) = \theta_i + \nu_i t + \sigma_i W(t), \tag{1}$$

where θ_i is the initial signal amplitude and ν_i is the drift of the process (see Figure 1b for a sample signal). The value of σ_i corresponds to standard deviation of the signal, which is assumed to be known and same across the population of generators, denoted by σ for all generators. The process $W(t)$ represents a standard Brownian motion with linear drift, where $W(0) = 0$. The increments $W(t+u) - W(t)$ for $u \geq 0$ are independent and identically distributed and follow a Normal distribution with mean 0 and variance u. We adopt the form (1) for degradation modeling as it is widely used in real-time condition monitoring (see Gebraeel et al. (2005)).

To model different degradation rates that correspond to the loading levels, we use the time transformation approach proposed by Doksum and Hóyland (1992) where the notion of *effective time* is used to scale the time under each stress condition. We extend the notion of effective time as follows.

Definition 1. *The effective time of generator i at time t is defined as $\tau_i(t) := \int_0^t L_i(t')dt'$, where the function $L_i(t)$ represents the load level of generator i.*

Here, the value of the load function $L_i(t)$ can be interpreted as a *load multiplier*. The value of this function is equal to 1 under nominal loading, which increases or decreases based on the dispatch decisions of the corresponding generator. In other words, the load function $L_i(t)$ is used to scale the signal $S_i(t)$ based on the loading condition. The scaled time is denoted by $\tau_i(t)$. In the case of nominal loading, $\tau_i(t) = t$. By using $\tau_i(t)$ in Equation 1, we can rescale the time period of different segments of the degradation signal by their respective load multiplier, thus allowing us to recreate a corresponding degradation signal with constant drift. It is noteworthy to mention that for the load-independent case $\tau_i(t)$ equals to t irrespective of the dispatch (loading) decisions, which may lead to inaccurate remaining life predictions.

2.2 Load-based remaining life estimation

Our underlying assumption is that different loading regimes will have the same effective time if their cumulative degradation is equivalent. In particular, we assume that a failure occurs when the degradation level of generator i, $S_i(t)$, reaches a predefined threshold value, Λ. In general settings, the parameters of the signal model (1) are unknown, and need to be estimated.

We assume that the unknown model parameters follow a prior distribution that can be estimated from historical data. The prior distribution represents the characteristics of the generator's population. The key assumption here is that the degradation process of the population exhibits a common functional trend. Specifically, we denote the prior distributions of θ_i and ν_i as $\pi_1(\theta_i)$ and $\pi_2(\nu_i)$. The prior distributions are assumed to follow a Normal distribution with mean μ_0 and variance σ_0^2, and mean μ_1 and variance σ_1^2, respectively. The random variables θ_i and ν_i are assumed to be mutually independent. The prior distribution will be updated using real-time signals observed from each generator using a Bayesian framework similar to the one proposed in Gebraeel et al. (2005). This allows the model to adapt to the unique degradation characteristics of each generator resulting in remaining life predictions that are driven by the generator's degradation process. To see this, assume that the signal levels at times $t_1^i, t_2^i, \cdots, t_k^i$ for every generator i are observed as $S_i(t_1^i), \ldots, S_i(t_k^i)$. As highlighted by Equation 1, we assume that the degradation signal follows a Brownian motion. Thus, we focus on modeling the increments of the signal, which we denote as $S_j^i = S_i(t_j^i) - S_i(t_{j-1}^i)$ for $j = 2, \cdots, k^i$, where $S_1^i = S_i(t_1^i)$. We assume that the future loading function of each generator i, $\{L_i(t) : t \geq 0\}$, is known a priori. By adopting a Bayesian updating approach (Proposition 2 in Gebraeel et al. (2005)) and utilizing the effective time notion (Definition 1), we can find the posterior distributions of θ_i and ν_i.

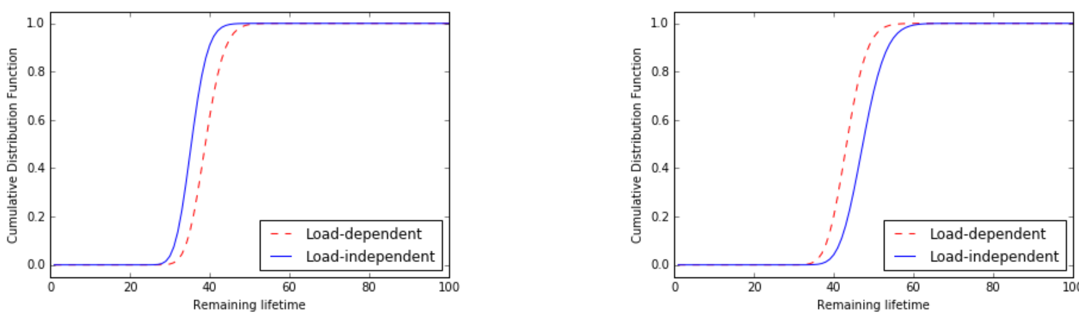
Proposition 1. *Given the observed signal increments, S_j^i, $j = 1, \cdots, k^i$, with parameters (ν_i, θ_i), and failure threshold Λ; for a load function, $\{L_i(t) : t \geq 0\}$, the posterior mean of the drift parameters ν_i, is given by,*

$$\mu_i' = \frac{(\sigma_1^2 \sum_{j=1}^{k^i} S_j^i + \mu_1 \sigma^2)(\sigma_0^2 + \sigma^2 t_1'^i) - \sigma_1^2 (S_1^i \sigma_0^2 + \mu_0 \sigma^2 t_1'^i)}{(\sigma_0^2 + \sigma^2 t_1'^i)(\sigma_1^2 t_k'^i + \sigma^2) - \sigma_0^2 \sigma_1^2 t_1'^i},$$

where $t_j'^i = \int_0^{t_j^i} L_i(t')dt'$, for $j = 1, \cdots, k$. Then, the corresponding remaining lifetime at t_k^i has an inverse Gaussian distribution $IG(w|\phi, \upsilon)$, where $w = \tau_i(t) = \int_0^t L_i(t_k^i + t')dt'$, $\phi = \frac{\Lambda - \sum_{j=1}^{k^i} S_j^i}{\mu_i'}$, and $\upsilon = \frac{(\Lambda - \sum_{j=1}^{k^i} S_j^i)^2}{\sigma_i^2}$.

Combining the Bayesian update procedure with the notion of effective time, we can compute updated load-dependent remaining lifetimes. We note that for the load-independent models, Proposition 1 can also be used for estimating the remaining lifetime distribution of signal i by replacing $t_j'^i$, with actual time t_j^i for $j = 1, \cdots, k$, and using $t_k^i + t$ in place of the effective time, $\tau_i(t)$.

To illustrate the difference in remaining life estimation between the load-dependent and load-independent approaches, we examine two specific cases. Consider a case where generator i is consistently operated under a high loading level until the k^{th} observation epoch, i.e., operating time, t_k^i. The effective time of this generator will be greater than the observed time, i.e., $t_k'^i > t_k^i$. Next, assume that the loading condition are switched and the generator will operate under a nominal load level for the rest of its lifetime. In a load-independent case, the remaining life distribution will be underestimated since it is based on an *inflated* drift parameter that assumes that the prevailing loading conditions remain the same. In contrast, the load-dependent model utilizes a drift value that has been adjusted based on the future loading level, nominal load. Figure 2a highlights the difference in the estimated cumulative distribution function of the remaining lifetime on a set of simulated signals using the two kinds of modeling approaches. The converse is also true. Figure 2b highlights the case where a generator operates at a less than nominal loading condition, i.e., $t_k'^i < t_k^i$, which overestimates the remaining lifetime once the future load function increases to nominal load.



(a) The case when $t_k'^i > t_k^i$.

(b) The case when $t_k'^i < t_k^i$.

Figure 2: CDFs of remaining lifetime for load-dependent and load-independent approaches.

2.3 Effective Time Estimation

In order to reflect the effects of operational decisions on degradation, one needs to accurately map the relationship between the load function and the decisions evaluated by the optimization model. In particular, the value of the load multiplier function $L_i(t)$ depends on the dispatch and maintenance decisions of generator i during period t while taking into account minimum and maximum production capacities. In this section, we discuss how we model the effective time based on the operational decisions and degradation uncertainty.

As mentioned before, we assume that load levels in each period are known, however drift values of the signals are uncertain. Our framework accommodates continuously varying load functions. Power generators, however, often operate under a fixed load level for a pre-specified period of time depending on their operating schedule. Consequently, it is reasonable to assume that load levels remain the same in between consecutive signal observations as the time of the signal observations correspond to the operational periods.

To estimate the effective time, we consider a set of historical degradation signals, namely $\mathcal{I}$, associated with generators that have been observed until their time of failure. In this context, our focus is on *soft failures* defined by unacceptable or alarming generator performance as opposed to *hard failures* that often result from catastrophic damage. We observe each signal $i \in \mathcal{I}$ until its failure time T_i, and consider the load function at the discrete time points from $0, 1, \cdots, T_i$. For $|J|$ levels of load, namely $L_1, L_2, \cdots, L_J$ and signal $i \in \mathcal{I}$, the mean estimate of the signal drift parameter, denoted by $\overline{\mu}_{i,j}$, can be estimated using the following expression:

$$\overline{\mu}_{i,j} = \frac{\sum_{t=1:L_i(t)=L_j}^{T_i} S_t^i}{\sum_{t=1}^{T_i} \mathbb{1}_{\{L_i(t)=L_j\}}}, \tag{2}$$

for $\sum_{t=1}^{T_i} \mathbb{1}_{\{L_i(t)=L_j\}} > 0$. In order to model the effective time, we estimate its value over each unit time. For this purpose, we normalize the mean estimates corresponding to each signal $i \in \mathcal{I}$ and load $j \in J$ pair, by dividing them the overall average signal drift, denoted by $\overline{\overline{\mu}}$. We denote the corresponding normalized estimates as $\overline{\mu}'_{i,j} = \overline{\mu}_{i,j}/\overline{\overline{\mu}}$.

We define the set of loading levels corresponding to the generator i as $L_{i,j} = p_i^{min} + (p_i^{max} - p_i^{min})(j-1)/(J-1)$ for $j = 2, \cdots, J-1$, and $L_{i,1} = p_i^{min}$, $L_{i,J} = p_i^{max}$, where p_i^{min} and p_i^{max} represent minimum and maximum production requirements of the corressponding generator. Using the $(L_{i,j}, \overline{\mu}'_{i,j})$ points, we develop a linear regression model to estimate the effective unit time, which we denote by $d_{i,t}$. We assume that generators have different capacities, and thus, we estimate an individual regression model for each generator. The resulting model for generator i can be represented as $d_{i,t} = \alpha'_i L'_i(t) + \beta'_i$, where $d_{i,t}$ is the estimated effective unit time at time t as a function of load level $L'_i(t)$, and α'_i and β'_i are the regression coefficients. We note that the load function $L'_i(t)$ is in terms of generation capacities. The relationship between the effective time $\tau_i(t)$ and $d_{i,t'}$ can be expressed as $\tau_i(t) = \sum_{t'=1}^{t} d_{i,t'}$.

To improve the practical relevance of our model, we assume that a generator does not degrade when it is not operating. Therefore, we integrate the commitment variable, $x_{i,t}$, into our regression model as follows; $d_{i,t} = \alpha'_i L'_i(t) + \beta'_i x_i$ where x_i is 1 when the generator operates and 0 otherwise. The regression models are foundational to characterizing the relationship between operational decisions and efficient maintenance scheduling of the generators.

3 Optimizing Maintenance and Operations

In this section, we formulate the load-dependent generator maintenance and operations scheduling problem as a decision-dependent stochastic program with cost and reliability perspectives. Given a fleet of generators, our aim is to obtain their maintenance and operations schedules while simultaneously minimizing maintenance and operations costs, and satisfying the system constraints under the load-dependency of generators' conditions. We consider a one-year planning horizon with monthly maintenance decisions, and daily operational schedules corresponding to commitment decisions, dispatch and demand curtailment amounts. We allow one maintenance per each generator during the planning horizon. Additionally, we consider a capacity limit on the number of ongoing maintenances. We note that a generator needs to be off if it is under maintenance. We also take into account operational level restrictions such as demand satisfaction, production capacities, and transmission line limits on the underlying power network.

Maintenance routines can be categorized into two groups. A preventive maintenance is conducted at the scheduled maintenance period, which costs C^p. Otherwise, a corrective maintenance is performed if a generator fails unexpectedly before its scheduled maintenance period with a cost, C^c. Corrective maintenance typically costs more and lasts longer compared to a scheduled maintenance. Thus, our aim is to identify cost effective and reliable maintenance and operations schedules that result in fewer number of unexpected failures with lower overall costs. To represent the trade-off between preventive and corrective maintenance, we adopt the *dynamic maintenance cost function* approach presented in Elwany and Gebraeel (2008); Yildirim et al. (2016a); Basciftci et al. (2018). The cost function uses the preventive and corrective maintenance costs coupled with the remaining life distribution of the generator to calculate the overall maintenance cost at future time epochs. We note that our framework enables updating the remaining lifetime estimations of the generators through newly acquired real-time degradation signals. This impacts the cost function, which is also dynamically revised to account for real-time changes in the degradation state of the generator. We extend the dynamic maintenance function definition to the load-dependent setting by integrating the effective time approach introduced in Section 2.1. We first define the decision variable $\tau_{i,t}$ as the effective age of generator i after t periods from the beginning of planning horizon of the optimization model. Next, the dynamic maintenance cost of generator i at time t with initial effective age $\tau_{i,0}$ can be expressed as follows:

$$C_{i,\tau_{i,0}}(\tau_{i,t}) = \frac{C^p \Pr(R_{i,\tau_{i,0}} > \tau_{i,t}) + C^c \Pr(R_{i,\tau_{i,0}} \leq \tau_{i,t})}{\int_0^{\tau_{i,t}} \Pr(R_{i,\tau_{i,0}} > z) dz \; + \tau_{i,0}}, \tag{3}$$

where $R_{i,\tau_{i,0}}$ is the remaining lifetime of generator i given the initial effective age $\tau_{i,0}$. We assume that the value of $\tau_{i,0}$ is known for every generator i at the beginning of planning.

Below is a summary of the sets, decision variables and parameters of the optimization model.

Sets:

- $\mathcal{B}$ Set of buses.
- $\mathcal{G}$ Set of generators.
- $\mathcal{L}$ Set of transmission lines.
- $\mathcal{S}$ Set of operational subperiods within a maintenance period.
- $\mathcal{T}$ Set of maintenance periods in the planning horizon.

Decision variables:

- $z_{i,t}$ 1 if generator i enters maintenance in maintenance period t, and 0 otherwise.
- γ_t Additional maintenance capacity added in maintenance period t.
- $\tau_{i,t}$ Effective age of generator i at time t.
- $x_{i,t,s}$ 1 if generator i is on in operational period s of maintenance period t, and 0 otherwise.
- $y_{i,t,s}$ Dispatch amount of generator i in operational period s of maintenance period t.
- $\psi_{b,t,s}$ Demand curtailed at bus b in operational period s of maintenance period t.

Parameters:

- C_{add} Per unit cost of maintenance capacity added.
- $V_{i,t,s}$ No-load cost of generator i in the operational period s of maintenance period t.
- $F_{i,t,s}$ Per unit dispatch cost of generator i in operational period s of maintenance period t.
- P_{DC} Per unit cost of demand curtailed.
- ξ Maintenance criticality coefficient.
- H Planning horizon length in maintenance periods.
- $\overline{M}$ Maximum number of ongoing maintenances.
- Y_p Duration of a preventive maintenance.
- ϵ Confidence level of the chance constraint.
- ρ Threshold on the number of generators to fail.
- $D_{b,t,s}$ Demand of bus b in operational period s of maintenance period t.
- p_i^{min} Minimum production requirement of generator i.

p_i^{max} Maximum production capacity of generator i.

f_{max}^l Flow capacity of line l.

a_l Shift factor vector for line l.

$M_{b,i}$ 1 if generator i is on bus b, and 0 otherwise.

The resulting load-dependent generator maintenance and operations scheduling problem can be formulated in (4) as follows:

$$\min \ \xi \left(\sum_{i\in\mathcal{G}} \sum_{t\in\mathcal{T}} C_{i,\tau_{i,0}}(\tau_{i,t}) z_{i,t} + \sum_{t\in\mathcal{T}} C_{add}\gamma_t \right) + \sum_{i\in\mathcal{G}} \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{S}} (V_{i,t,s} x_{i,t,s} + F_{i,t,s} y_{i,t,s}) + \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{S}} \sum_{b\in\mathcal{B}} P_{DC}\psi_{b,t,s} \tag{4a}$$

$$\text{s.t.} \ \tau_{i,t} = \sum_{t'=1}^{t} \sum_{s\in\mathcal{S}} (\alpha_i y_{i,t',s} + \beta_i x_{i,t',s}) \qquad i\in\mathcal{G}, t\in\mathcal{T} \tag{4b}$$

$$\Pr\left(\sum_{i\in\mathcal{G}} \sum_{t\in\mathcal{T}} \zeta_{i,t}(\tau_{i,t}) z_{i,t} \le \rho \right) \ge 1-\epsilon \tag{4c}$$

$$\sum_{e=0}^{Y_p-1} z_{i,t-e} \le \overline{M} + \gamma_t \qquad t\in\mathcal{T} \tag{4d}$$

$$\sum_{t\in\mathcal{T}} z_{i,t} = 1 \qquad i\in\mathcal{G} \tag{4e}$$

$$x_{i,t,s} \le 1 - \sum_{e=0}^{Y_p-1} z_{i,t-e} \qquad i\in\mathcal{G}, t\in\mathcal{T}, s\in\mathcal{S} \tag{4f}$$

$$\sum_{i\in\mathcal{G}} y_{i,t,s} + \sum_{b\in\mathcal{B}} \psi_{b,t,s} = \sum_{b\in\mathcal{B}} D_{b,t,s} \qquad t\in\mathcal{T}, s\in\mathcal{S} \tag{4g}$$

$$p_i^{min} x_{i,t,s} \le y_{i,t,s} \le p_i^{max} x_{i,t,s} \qquad i\in\mathcal{G}, t\in\mathcal{T}, s\in\mathcal{S} \tag{4h}$$

$$\left| \sum_{b\in\mathcal{B}} a_{l,b} \left(\sum_{i\in\mathcal{G}} M_{b,i} y_{i,t,s} + \psi_{b,t,s} - D_{b,t,s} \right) \right| \le f_l^{max} \qquad t\in\mathcal{T}, s\in\mathcal{S}, l\in\mathcal{L} \tag{4i}$$

$$z_{i,t}, x_{i,t,s} \in \{0,1\}, \gamma_t, y_{i,t,s} \ge 0, D_{b,t,s} \ge \psi_{b,t,s} \ge 0 \quad i\in\mathcal{G}, t\in\mathcal{T}, s\in\mathcal{S}, b\in\mathcal{B}. \tag{4j}$$

Objective function (4a) minimizes total maintenance and operations cost of a fleet of generators. The first part of the objective represents the maintenance cost, in particular the dynamic maintenance cost (3) and additional labor costs. To approximate the maintenance cost function, we propose a piecewise linearization approach, which is described in detail in Section 3.2. The remaining part of the objective corresponds to the operational costs including the costs of commitment, dispatch, and demand curtailment. The cost of maintenance is adjusted with respect to the cost of operations by the maintenance criticality coefficient, ξ. Selection of ξ values enables examining the importance of the maintenance and operations costs on the resulting schedules.

Constraint (4b) represents the effective age formulation. The modeling of this constraint and the derivation of the corresponding parameters (α', β') are described in detail in Section 2.3. In order to ensure that the variable $\tau_{i,t}$ represents the effective age of the generator i in terms of maintenance periods, the parameters (α_i, β_i) for each generator i are taken as $\alpha_i = \alpha'_i/|\mathcal{S}|$, and $\beta_i = \beta'_i/|\mathcal{S}|$. This change of parameters helps in establishing the time transformation from operational periods (i.e. days) to maintenance periods (i.e. months).

The chance constraint (4c) aims to restrict the number of generators that fail before their scheduled maintenance with a threshold ρ with high probability $1-\epsilon$. This constraint leverages sensor information through the random variable ζ. The Bernoulli random variable $\zeta_{i,t}$ is 1 if $\tau_{i,t} \geq R_{i,\tau_{i,0}}$ and 0 otherwise. This constraint formulates a decision-dependent uncertainty, as the failure probabilities depend on $\tau_{i,t}$, which is related with the dispatch and commitment decisions through constraint (4b). As the chance constraint is computationally intractable, we develop a combination of safe approximation and piecewise linearization approaches for its representation in Section 3.1 and Section 3.2, respectively.

Constraint (4d) guarentees that there is at most $\overline{M} + \gamma_t$ maintenances in each period t. Thus, the maintenance capacity of the system can be violated in return for its penalty in the objective. Constraint (4e) ensures that each generator enters maintenance once through the planning horizon, which is a common assumption in generator maintenance scheduling in power systems literature (see e.g. Conejo et al. (2005), Wang et al. (2016)). Constraint (4f) enforces the generators to be off if they are under maintenance.

The remaining constraints in (4) represent the operational level restrictions. In particular, constraint (4g) ensures that total demand is satisfied with production and demand curtailment. Constraint (4h) guarentees that generators produce within their production limits, and constraint (4i) enforces the transmission line limits by considering the DC approximation (see Cain et al. (2012)) for modeling the power flow.

3.1 Safe approximation of the chance constraint

The proposed chance constraint (4c) poses computational challenges, as it is intractable to represent and considers decision-dependent uncertainty. For this purpose, we present alternative ways for reexpressing this constraint. We utilize a deterministic safe approximation of the chance constraint as follows:

Proposition 2. *The deterministic constraint*

$$\sum_{i\in\mathcal{G}}\sum_{t\in\mathcal{T}}\mathbf{E}[\zeta_{i,t}(\tau_{i,t})]\, z_{i,t} \leq \max\left(\rho\,\epsilon, \max_{\delta>0}\left[\frac{((\epsilon\, e^{\delta\rho})^{1/|\mathcal{G}|}-1)|\mathcal{G}|}{e^{\delta}-1}\right]\right) = \rho^* \tag{5}$$

is a safe approximation of (4c)*, i.e. any* $z \in \{0,1\}^{|\mathcal{T}|\times|\mathcal{G}|}$ *satisfying* (5)*, satisfies* (4c)*.*

We note that Proposition 2 is an extension of Proposition 1 in Basciftci et al. (2018), in which the random variable $\zeta_{i,t}$ is independent of the effective age of the generator $\tau_{i,t}$.

The term $\mathbf{E}[\zeta_{i,t}(\tau_{i,t})]$ in Proposition 2 can be expressed as $\mathbf{E}[\zeta_{i,t}(\tau_{i,t})] = \Pr(R_{i,\tau_{i,0}} \leq \tau_{i,t})$, using the definition of the Bernoulli random variable ζ. To represent this decision-dependent uncertainty, we define an auxiliary decision variable $P_{i,t} := \mathbf{E}[\zeta_{i,t}(\tau_{i,t})]$. Considering $\bar{P}_{i,t}$ as an upper bound on $P_{i,t}$, and utilizing $0 \leq P_{i,t} \leq \bar{P}_{i,t} \leq 1$, we can linearize the term $v_{i,t} := P_{i,t}\, z_{i,t}$. The safe approximation of the chance constraint (4c) is represented in the form in (6).

$$\sum_{i\in\mathcal{G}}\sum_{t\in\mathcal{T}} v_{i,t} \leq \rho^* \tag{6a}$$

$$0 \leq v_{i,t} \leq P_{i,t},\;\; P_{i,t} - (1-z_{i,t})\bar{P}_{i,t} \leq v_{i,t} \leq \bar{P}_{i,t} z_{i,t} \tag{6b}$$

Similarly, the objective function (4a) includes nonlinear terms. To handle this issue, we linearize $C_{i,\tau_{i,0}}(\tau_{i,t})z_{i,t}$. Let $\theta_{i,t}$ be $C_{i,\tau_{i,0}}(\tau_{i,t})$. Then, we define $w_{i,t} := \theta_{i,t}z_{i,t}$. Since the cost of corrective maintenance is an upper bound on the dynamic maintenance cost function, we observe that $0 \leq \theta_{i,t} \leq \bar{\theta}_{i,t} \leq C^c$, where $\bar{\theta}_{i,t}$ is an upper bound on $\theta_{i,t}$. Thus, we linearize $w_{i,t}$ as follows:

$$0 \leq w_{i,t} \leq \theta_{i,t},\;\; \theta_{i,t} - (1-z_{i,t})\bar{\theta}_{i,t} \leq w_{i,t} \leq \bar{\theta}_{i,t} z_{i,t} \tag{7}$$

3.2 Piecewise linearization

We note that for a generator i, its probability of failure by time t, $P_{i,t}$, and maintenance cost at time t, $\theta_{i,t}$, depend nonlinearly on its effective age $\tau_{i,t}$. To accurately capture these nonlinear relationships, we propose linearization procedures for representing $P_{i,t}$ and $\theta_{i,t}$ as functions of $\tau_{i,t}$. For this purpose, we examine the maintenance cost function and the remaining lifetime distribution of each generator under specific breakpoints, namely $d_i^0, d_i^1, \cdots, d_i^K$ for every generator i. Then, we find the corresponding failure probabilities and dynamic maintenance cost function values evaluated at the breakpoints as $P_i^k = \Pr(R_{i,\tau_{i,0}} \leq d_i^k)$ and $\theta_i^k = C_{i,\tau_{i,0}}(d_i^k)$ for $k = 1, \cdots, K$, respectively. We illustrate the nonlinearity of the maintenance cost function and its associated piecewise approximation on a sample signal in Figure 3. Since we have monthly maintenance decisions, we utilize monthly breakpoints as shown. As remaining lifetime distribution and maintenance cost functions are not convex, we need *special ordered sets of type 2 (SOS2)* constraints in the piecewise linearization, which are a form of disjunctive constraints, see Vielma and Nemhauser (2008). To formulate these constraints, we consider two formulations studied in Vielma et al. (2010) with linearly or logarithmically many extra binary variables in the number of breakpoints and constraints. We refer to the first case as *linear* formulation, and second one as *log* formulation. Our preliminary computational results illustrate the significant computational advantage of the log formulation over the linear formulation (see Table 4). Therefore, we focus on the log formulation in the remainder of the paper.

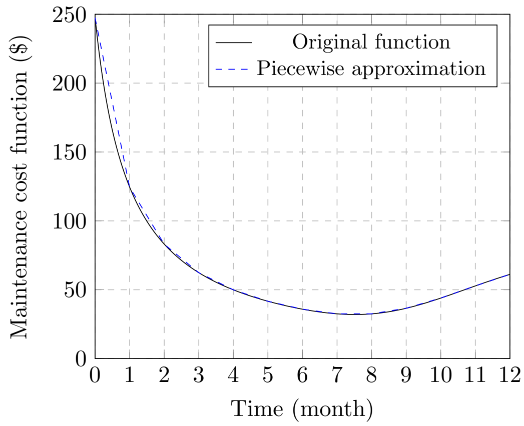


Figure 3: Piecewise linearization of maintenance cost function.

The corresponding model can be represented by defining the additional variables $\lambda_{i,t}^k \geq 0$ and $\eta_{i,t}^m \in \{0,1\}$, where $\sum_{k=0}^{K} \lambda_{i,t}^k = 1$ for all $i \in \mathcal{G}$, $t \in \mathcal{T}$, $k = 0, 1, \cdots, K$, $m \in \mathcal{M}$. The variable $\eta_{i,t}^m$ depends on $\lambda_{i,t}^k$ as described in Theorem 1.

Theorem 1. *(Theorem 1 in Vielma and Nemhauser (2008)) Let* $B : \{1, \cdots, K\} \rightarrow \{0,1\}^{\lceil \log_2 K \rceil}$ *be an SOS2 compatible function, i.e. a function that enforces SOS2 constraints on* $\{\lambda_{i,t}^k\}_{k=0}^{K} \in \mathbb{R}_{+}^{K+1}$ *if for all* $l \in \{1, \cdots, K-1\}$ *the vectors* $B(l)$ *and* $B(l+1)$ *differ in at most one component. Then the following inequalities are valid for SOS2 constraints:*

$$\sum_{k \in K^{+}(m,B)} \lambda_{i,t}^k \leq \eta_{i,t}^m \quad m \in \mathcal{M},\ i \in \mathcal{G},\ t \in \mathcal{T} \tag{8a}$$

$$\sum_{k \in K^{0}(m,B)} \lambda_{i,t}^k \leq (1 - \eta_{i,t}^m) \quad m \in \mathcal{M},\ i \in \mathcal{G},\ t \in \mathcal{T} \tag{8b}$$

where $\mathcal{M} = \{1, \cdots, \lceil \log_2 K \rceil\}$, $K^{+}(m,B) = \{j \in J : \forall i \in I(j)\ m \in \sigma(B(i))\}$, *and* $K^{0}(m,B) =$

$\{j \in J : \forall i \in I(j) \;\; m \notin \sigma(B(i))\}$. *The function* $\sigma(r)$ *represents the support of vector* r*, which corresponds to the set of indices of* r *such that* $r_i \neq 0$*, and the sets* $I = \{1, \cdots, K\}$*,* $J = \{0, 1, \cdots, K\}$*,* $S_i = \{i-1, i\}$ *for all* $i \in I$*, and* $I(j) = \{i \in I : j \in S_i\}$ *for all* $j \in J$*.*

We note that we use Gray code (Wilf, 1989) as the SOS2 compatible function in our formulation, which is used in binary numeral systems to order numbers in such a way that a pair of successive numbers are only different in one binary digit. The remaining constraints, in addition to (8), for the piecewise linearization can be expressed as follows:

$$\sum_{k=0}^{K} \lambda_{i,t}^{k} \theta_i^k = \theta_{i,t} \qquad i \in \mathcal{G},\; t \in \mathcal{T} \tag{9a}$$

$$\sum_{k=0}^{K} \lambda_{i,t}^{k} P_i^k = P_{i,t} \qquad i \in \mathcal{G},\; t \in \mathcal{T} \tag{9b}$$

$$\sum_{k=0}^{K} \lambda_{i,t}^{k} d_i^k = \tau_{i,t} \qquad i \in \mathcal{G},\; t \in \mathcal{T} \tag{9c}$$

$$\sum_{k=0}^{K} \lambda_{i,t}^{k} = 1 \qquad i \in \mathcal{G},\; t \in \mathcal{T} \tag{9d}$$

$$\lambda_{i,t}^{k} \geq 0 \quad k \in \{0, 1, \cdots, K\}, \quad \eta_{i,t}^{m} \in \{0,1\} \qquad m \in \mathcal{M},\; i \in \mathcal{G},\; t \in \mathcal{T} \tag{9e}$$

3.3 Formulation enhancements

We improve the resulting formulation by benefiting from the underlying structure of the problem. The constraint (4e) ensures that each generator enters maintenance once during the planning horizon. Therefore, we propose an alternative effective time definition by only considering the values of $\tau_{i,t}$ variables at the time the generators enter maintenance. This information is sufficient for our formulation to compute the maintenance cost function, and the failure probabilities in the chance constraint. Let $\tau'_{i,t} = \tau_{i,t}\, z_{i,t}$. To incorporate this variable in formulation (8) and (9), we revise the constraints (9c) and (9d) as follows:

$$\sum_{k=0}^{K} \lambda_{i,t}^{k} d_i^k = \tau'_{i,t} \quad i \in \mathcal{G},\; t \in \mathcal{T} \tag{10a}$$

$$\sum_{k=0}^{K} \lambda_{i,t}^{k} = z_{i,t} \quad i \in \mathcal{G},\; t \in \mathcal{T} \tag{10b}$$

$$\eta_{i,t}^{m} \leq z_{i,t} \quad m \in \mathcal{M},\; i \in \mathcal{G},\; t \in \mathcal{T} \tag{10c}$$

When $z_{i,t}$ is 0, then $\lambda_{i,t}^{k}$ and $\eta_{i,t}^{m}$ values are set to 0 for all $k = 0, \cdots, K$ and $m \in \mathcal{M}$, because of (10b) and (10c), respectively. Thus, $\tau'_{i,t}$ becomes 0 as desired.

This approach provides a significant computational advantage by eliminating the consideration of $\tau_{i,t}$ values when $z_{i,t} = 0$. Similar to the previous linearizations, (6) and (7), the constraint set $\tau'_{i,t} = \tau_{i,t}\, z_{i,t}$ is linearized, by defining the upper bound value of $\tau'_{i,t}$ as $\overline{\tau}_{i,t}$.

Combining the above, the resulting mathematical problem for the load-dependent maintenance and optimization scheduling (4) is reformulated in (11) as a mixed-integer linear program. We remove the decision variable $\tau_{i,t}$ from the model as it is no longer needed explicitly.

$$\min\ \xi \left(\sum_{i\in\mathcal{G}} \sum_{t\in\mathcal{T}} w_{i,t} + \sum_{t\in\mathcal{T}} C_{add}\gamma_t \right) + \sum_{i\in\mathcal{G}} \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{S}} (V_{i,t,s}x_{i,t,s} + F_{i,t,s}y_{i,t,s}) + \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{S}} \sum_{b\in\mathcal{B}} P_{DC}\psi_{b,t,s} \tag{11a}$$

$$\text{s.t. } (4\text{d}) - (4\text{j}), (6) - (8), (9\text{a}), (9\text{b}), (10)$$

$$\tau'_{i,t} \le \overline{\tau}_{i,t} z_{i,t} \qquad i\in\mathcal{G}, t\in\mathcal{T} \tag{11b}$$

$$\tau'_{i,t} \le \sum_{t'=1}^{t} \sum_{s\in\mathcal{S}} (\alpha_i y_{i,t',s} + \beta_i x_{i,t',s}) \qquad i\in\mathcal{G}, t\in\mathcal{T} \tag{11c}$$

$$\tau'_{i,t} \ge \sum_{t'=1}^{t} \sum_{s\in\mathcal{S}} (\alpha_i y_{i,t',s} + \beta_i x_{i,t',s}) - (1 - z_{i,t})\overline{\tau}_{i,t} \qquad i\in\mathcal{G}, t\in\mathcal{T} \tag{11d}$$

$$\tau'_{i,t} \ge 0, \lambda^k_{i,t} \ge 0,\ k = 0, 1, \cdots, K,\ i\in\mathcal{G},\ t\in\mathcal{T},\ \eta^m_{i,t} \in \{0,1\}\ m\in\mathcal{M},\ i\in\mathcal{G},\ t\in\mathcal{T} \tag{11e}$$

In order to improve the upper bound values used in the linearization, we consider the effect of the load decisions on the data-driven degration equivalent time model. In any time t, the corresponding effective time for generator i, i.e. $\tau_{i,t}$, can be at most $\sum_{t'=1}^{t} \sum_{s\in\mathcal{S}} (\alpha_i p_i^{max} + \beta_i)$, as ensured by the constraint (11c). Thus, we can select

$$\overline{\tau}_{i,t} := t\,|\mathcal{S}|(\alpha_i p_i^{max} + \beta_i). \tag{12}$$

Similarly, we can identify upper bounds for the failure probabilities $P_{i,t}$ and the dynamic maintenance cost function $\theta_{i,t}$ for generator i at time t. Since $P_{i,t} = \Pr(R_{i,\tau_{i,0}} \le \tau_{i,t})$ is monotonically nondecreasing with respect to degradation amount, we can take its upper bound value as $\bar{P}_{i,t} = \Pr(R_{i,\tau_{i,0}} \le \overline{\tau}_{i,t})$. Finally, we can obtain an upper bound value for the dynamic maintenance cost function $\theta_{i,t}$ as $\bar{\theta}_{i,t} = \max_{0\le t' \le \bar{\tau}_{i,t}} C_{i,\tau_{i,0}}(t')$. We note that we are not able to simply select $C_{i,\tau_{i,0}}(\overline{\tau}_{i,t})$ as the upper bound value, since the cost function is not necessarily monotonic with respect to effective time.

4 Computational results

In this section, we provide a comprehensive framework to illustrate the effectiveness of our approach. We first discuss the experimental setup to estimate remaining lifetime and effective time of each generator in Section 4.1 and Section 4.2, respectively. To evaluate the performances of different maintenance and operations schedules, we develop a decision-dependent simulation procedure in Section 4.3. We provide our computational experiments in Section 4.4 by studying various instances under different congestion levels and reliability considerations. Finally, we illustrate the computational gains of the proposed algorithmic enhancements for solving the optimization model in Section 4.5. We note that the flowchart of our computational framework is presented in Figure 4.

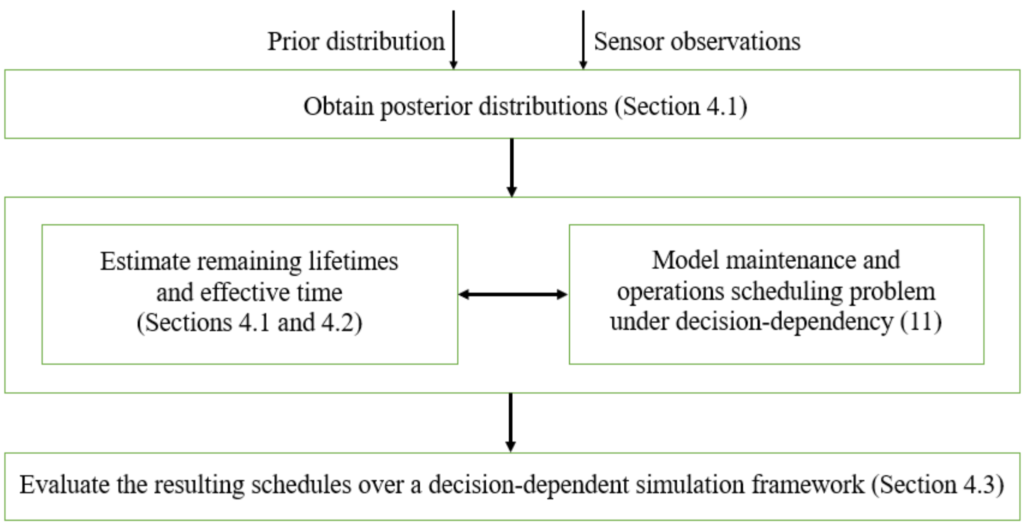


Figure 4: Flowchart of the computational framework.

4.1 Determining prior distribution and remaining lifetime estimation

To estimate prior distributions corresponding to signal characteristics in (1), we first construct a set of 100 signals under different load levels. These signals mimic the degradation process of a rotating bearing, and follow the form (1) with $\theta_i \sim N(20, 3^2)$, $\nu_i \sim N(2.5, 0.2^2)$ and $\sigma = 3.5$. We examine the signal values at discrete time points, and assume that load level remains constant between consecutive observations. We observe the signals until a failure threshold Λ, which is taken as 150. Time of failure of each signal i is denoted as T_i. As before, we represent the differences in the observations of each signal i as S_k^i, where $S_1^i = S_i(0)$, and $S_k^i = S_i(k) - S_i(k-1)$ for $k = 2, \cdots, T_i$. Similarly, $L_{i,j}$ corresponds to the load level of signal i in period j. We assume that the variance of the stochastic parameters are known. Thus, we only need to estimate the mean of the prior distributions of the stochastic model parameters θ and ν, which are μ_0 and μ_1 respectively. As the initial amplitude of each signal i, i.e. S_1^i, corresponds to θ_i values in the form (1), we compute the mean estimate of $\pi_1(\theta)$ by averaging these values over the set of signals. To estimate the mean of the prior distribution of ν, we find the mean estimates, $\hat{\mu_i}$ corresponding to each signal i, $i = 1, \cdots, 100$. For finding these estimates for the load-dependent models, we adopt the time transformation concept discussed in Section 2.2. In particular, we can estimate $\hat{\mu_i}$ as

$$\hat{\mu_i} = \frac{\sum_{j=1}^{T_i} S_j^i - S_1^i}{\sum_{j=1}^{T_i} L_{i,j}}, \tag{13}$$

for $i = 1, \cdots, 100$, where the denominator corresponds to an effective time estimate at failure. By averaging these estimates, we obtain the mean estimate of the prior distribution $\pi_2(\nu)$. As the load-independent models neglect the load decisions in determining degradation amount, they consider a different estimate for the prior distribution $\pi_2(\nu)$. Specifically, mean estimate of each signal can be computed as $(\sum_{j=1}^{T_i} S_j^i - S_1^i)/T_i$ by only considering the operational time until failure. Consequently, we obtain the prior estimate for the load-independent model by averaging these values over 100 signals.

In order to represent the signal characteristics specific to each generator, we combine the prior distributions with sensor information. For this purpose, we assign signals to each generator $i \in \mathcal{I}$, which are different than the 100 signals used in prior estimation. These signals are partially degraded at the beginning of the planning horizon with a random initial age. Using these observations until the time of planning, we obtain the posterior distribution of the unknown parameters as discussed in Proposition 1. After obtaining these component specific estimates, we identify the remaining lifetime distribution corresponding to each generator to be used in the optimization model.

4.2 Effective time estimation

We utilize the effective time estimation procedure described in Section 2.3 using the signals for estimating prior distribution parameters. We consider 3 levels of load, i.e. L_1, L_2, L_3, which corresponds to values 0.5, 1.0, 1.5 respectively. Then, for each signal and load level, we find the estimates using (2). Therefore, we obtain the set of points $(L_j, \overline{\mu}_{i,j})$ for each load level $j \in \{1, 2, 3\}$ and signal $i \in \{1, \cdots, 100\}$. These points for the given signal set are illustrated in Figure 5, which shows the variability in the signals under each load level. We note that the overall drift average over the signals, $\overline{\overline{\mu}}$, is computed as 2.49.

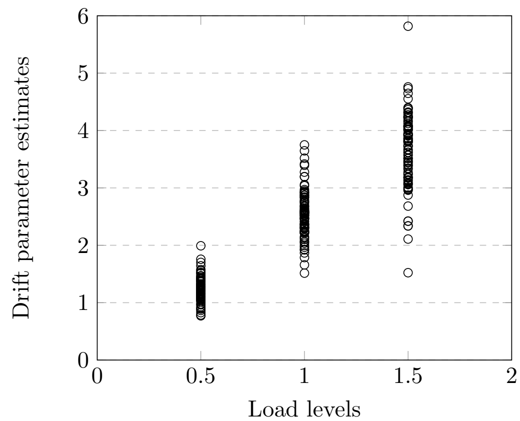


Figure 5: Signal variability under each load level.

For every generator i, we consider the values p_i^{min}, $(p_i^{min} + p_i^{max})/2$, and p_i^{max} corresponding to the load levels L_1, L_2, L_3, respectively. Next, we rescale the drift parameter estimates by dividing them to the population mean estimate $\overline{\overline{\mu}}$. Consequently, the rescaled vertical axis represents the unit effective time. By applying regression analysis specific to each generator i, we find the estimates α_i and β_i to be used in the optimization model (11).

We remind the reader that load-independent models disregard the notion of effective time based on the production decisions. To capture this approach, we replace constraint (4b) with

$$\tau_{i,t} = \sum_{t'=1}^{t} \sum_{s \in \mathcal{S}} \frac{1}{|\mathcal{S}|} x_{i,t',s} \quad i \in \mathcal{G}, t \in \mathcal{T}, \tag{14}$$

which gives the operational age of the generator. Then, we proceed with the same formulation methodology as in the load-dependent case in Section 3.2 and Section 3.3 to linearize and represent the variable $\tau'_{i,t}$. Consequently, effective time in these models are only based on the operational age of the generators. By coupling this condition with the remaining lifetime estimations specific to the load-independent models, we can model the maintenance and operations scheduling problem under solely operational time-based degradation.

4.3 Simulation and solution evaluation

In order to compare the performances of different maintenance and operations schedules, we propose a decision-dependent simulation procedure. In each period, we simulate the degradation process of each generator by creating signals based on the signal characteristics, and the dispatch and commitment decisions. For simulation purposes, we assume that the true distribution of θ_i and μ_i values in the functional form (1) are known for each generator $i \in G$. For representing effective time, we use α_i, β_i values found in Section 4.2.

Algorithm 1 in Appendix A describes the proposed methodology to evaluate a given maintenance and operations schedule in detail. We start the simulation procedure by considering the last observed signal amplitude of each generator at the beginning of planning. Since we observe each signal i until time t_k^i, we represent the last amplitude as $\sum_{j=1}^{k^i} S_j^i$ as discussed in Section 2.2. Then, we simulate each generator's corresponding signal under the given operations schedule. At the end of each maintenance period, we check the condition of the generators by observing their signal amplitudes. If the period t is the scheduled maintenance time of generator i, i.e. $z_{i,t} = 1$, and the generator has not failed previously, then generator enters preventive maintenance and remains closed for Y_p periods. If the signal amplitude of generator i, namely Amp_i, is greater than the failure threshold Λ, then generator i fails. It enters corrective maintenance immediately and stays closed for Y_c periods. As failures are unexpected, corrective maintenance requires more resources than a scheduled maintenance, i.e. $Y_c > Y_p$. After a maintenance ends, a new signal is assigned to that generator to represent its degradation process in the remainder of the planning horizon. Since components start degrading after their first phase ends, the degradation process starts after maintenance is completed, and first phase is over.

When a generator fails unexpectedly, it will not be able to produce in the upcoming Y_c periods. This unexpected loss in the production needs to be explicitly taken into account while evaluating the maintenance schedule. For this purpose, we consider these types of losses in production due to failures as demand curtailment while computing the operational cost in simulation.

4.4 Computational Experiments

In this section, we present a comprehensive computational study by comparing the performance of the solutions from load-dependent and load-independent models. We evaluate these solutions using the simulation procedure described in Section 4.3. We provide our computational results on 39-bus New-England Power System (Athay et al., 1979), and 118-bus instances (Blumsack, 2006). An overview of the instances is provided in Table 1, and further details of the power system configurations are discussed in the aforementioned papers and references therein. We implement the proposed model with enhancements (11) in Python using Gurobi 7.5.2 as the solver on an Intel i5-3470T 2.90 GHz machine with 8 GB RAM.

	# Buses	# Lines	# Generators	Total capacity (MWh)
39-bus	39	46	10	8840.4
118-bus	118	186	19	5859.2

Table 1: Overview of the Instances.

We study a one-year maintenance plan with monthly maintenance and daily operational decisions. For the chance constraint (4c), we set ρ as $\lfloor |\mathcal{G}|/3 \rfloor$ with $\epsilon = 0.05$ or 0.10. This implies that at most one third of the generators enters corrective maintenance due to a failure with a probability of at least $(1 - \epsilon)$. The safe approximation discussed in Proposition 2 is used to represent the chance constraint. We set cost of preventive maintenance C^p = \$100.000, and corrective maintenance C^c = \$400.000. These cost values are used in both dynamic maintenance cost function calculation in (3), and in the simulation for evaluating maintenance costs. To observe the performance of the proposed approach under various signal characteristics, we generate a partially degraded set of signals following the procedure in Section 4.1. Then, we randomly assign these signals to the generators and repeat each experiment 5 times with different set of signals. For each setting, we report the average results of these 5 macro-replications.

We evaluate three modeling approaches with respect to their remaining lifetime estimation procedures and optimization formulations, namely i) load-dependent, ii) load-independent, and iii) reliability-based. Load-dependent refers to the proposed approach of the paper to represent the decision-dependent degradation in maintenance and operations scheduling. Load-independent and reliability-based approaches consider an operational age-based degradation modeling as represented in (14) in the optimization model (11), whereas they differ in their remaining lifetime estimation procedures. Load-independent approach adopts its estimation procedure described in Section 2.2. For the reliability-based case, we derive the lifetime distributions by first fitting an inverse Gaussian distribution to a given set of failure points of the signals used in prior estimation, and then conditioning to the initial ages of the generators.

As demand level of the system plays an important role in the maintenance and operations decisions and the degradation amount of the generators, we study the instances under two congestion levels. Table 2 and Table 3 correspond to the results under high and low system congestions, respectively. In particular, low and high congestion correspond to the cases where the average daily demand of the system over a yearly planning horizon is adjusted to be 40% and 70% of the system capacity. The columns '# of failures', 'MC' and 'TC' represent the average number of failures, maintenance and total cost (sum of maintenance and operations costs) in the evaluated simulation procedure for a given solution. Each instance is studied under two different

	ϵ	Type	# of Failures	MC ($M)	Gain (%)	TC ($M)	Gain (%)
39-bus	0.05	LD	0.42	1.13		113.30	
		LI	1.31	1.39	18.74	129.62	12.59
		RB	0.63	1.19	4.94	117.62	3.68
	0.10	LD	0.26	1.08		111.14	
		LI	1.18	1.35	20.24	124.84	10.97
		RB	1.04	1.32	17.91	124.13	10.47
118-bus	0.05	LD	0.66	2.10		74.00	
		LI	2.11	2.53	16.97	80.79	8.41
		RB	1.64	2.40	12.25	80.20	7.73
	0.10	LD	0.60	2.08		72.24	
		LI	2.14	2.54	18.09	81.42	11.27
		RB	1.63	2.39	12.89	80.58	10.35

Table 2: Solution Evaluation under High Congestion.

reliability levels of chance constraint, which is adjusted by the parameter ϵ. The abbreviations 'LD', 'LI', 'RB' are used for load-dependent, load-independent and reliability-based approaches, respectively.

	ϵ	Type	# of Failures	MC ($M)	Gain (%)	TC ($M)	Gain (%)
39-bus	0.05	LD	0.04	1.01		53.02	
		LI	0.60	1.18	14.12	59.82	11.36
		RB	0.10	1.03	1.79	53.83	1.51
	0.10	LD	0.06	1.02		53.14	
		LI	0.65	1.20	14.82	60.08	11.54
		RB	0.12	1.04	1.79	53.75	1.13
118-bus	0.05	LD	0.08	1.93		37.30	
		LI	0.94	2.18	11.58	40.70	8.35
		RB	0.49	2.05	5.87	38.33	2.68
	0.10	LD	0.05	1.92		37.31	
		LI	1.06	2.22	13.49	41.46	10.01
		RB	0.51	2.05	6.55	38.31	2.62

Table 3: Solution Evaluation under Low Congestion.

As the proposed load-dependent approach captures the effect of operational decisions on degradation modeling within the optimization model, it performs better in terms of number of failures and maintenance cost in comparison to load-independent and reliability-based approaches. We observe 5-20% and 2-15% maintenance cost savings of load-dependent approach in high and low system congestions, respectively, compared to the previously studied methods in the literature.

When a generator fails unexpectedly, there is an unplanned loss in production capacity. This disruption in the operational schedule is penalized with demand curtailment cost in the solution evaluation. Our analyses highlight significant cost savings in total cost in the order of 3-13% and 1-11% for high and low congestion cases, by adopting the load-dependent approach. Furthermore, when systems are under high congestion, we observe more failures in all instances. This happens since high demand levels initiate higher levels of production resulting in faster degradation.

We emphasize that load-independent and reliability-based approaches are insufficient in truly representing the dependency between the degradation modeling and optimization framework. Nevertheless, as we compare the two approaches, we observe that reliability-based schedules perform better in solution evaluation compared to load-independent schedules. Reliability-based remaining lifetime estimations consider more variance in data by fitting a lifetime distribution based on failure points, whereas load-independent estimations are tailored to unit specific observations. Consequently, maintenance cost function and remaining lifetime estimations of reliability-based approach do not change much between different maintenance decisions, compared to load-independent models. Thus, the resulting optimization model becomes less sensitive to the choice of maintenance and operations schedules.

We note that as ϵ value gets larger, the chance constraint becomes less restrictive. Although the choice of ϵ does not necessarily affect the performance of the resulting schedule in high congestion case, average number of failures decreases when $\epsilon = 0.05$ in low congestion setting. Overall, the results demonstrate that load-dependent solutions outperform load-independent and reliability-based solutions with a smaller number of failures, and lower maintenance and operational costs in all the settings considered.

4.5 Computational efficiency

In this section, we illustrate the computational efficiency with respect to different forms of enhancements. We first examine the computational advantage of the selected piecewise linearization procedure, used for linearizing the objective and the safe approximation of the chance constraint. Specifically, we compare linear and log formulations described in Section 3.2. Secondly, we illustrate the performance of the proposed formulation enhancements in Section 3.3. Lastly, we demonstrate the effect of a *priority branching* method. More specifically, as maintenance decisions play an important role in determining the effective time, we put a special emphasis on those variables while solving the problem. For this purpose, we employ a priority branching method, which is used in optimization for directing the branch-and-bound procedure. By prioritizing the variables corresponding to the maintenance decisions, z, over the commitment decisions x, the respective branch-and-bound tree prefers branching on the maintenance variables.

We demonstrate the results on sample 39-bus instances under high congestion and $\epsilon = 0.05$ in Table 4. We report the average run time results over 5 macro-replications. We note that the linear formulation is not able to converge in 10000 seconds in all replications. By log formulation, we refer to the proposed integrated maintenance and operations scheduling model without the formulation enhancements introduced in Section 3.3.

	Run time	Speed-up
Linear formulation	>10000.00	
Log formulation	1189.89	×8.40
Log formulation with enhancements (11)	601.53	×16.62
Log formulation with enhancements (11) + priority branching	576.41	×17.35

Table 4: Run time (seconds) comparison on sample instances.

The results show that each enhancement significantly contributes to the run time performance. Priority branching coupled with the formulation enhancements gives the best results with an overall speedup of more than 17 times, demonstrating the significant computational gains of the proposed improvements.

5 Conclusion

In this study, we present a comprehensive framework for effectively solving condition-based maintenance and operations scheduling problem of a fleet of generators under load-dependency. We propose a data-driven degradation modeling framework to capture the endogenous effect of the operational decisions. First, we present a sensor-driven remaining lifetime estimation procedure under time-varying load decisions. We also develop an estimation method to capture the effect of the load decisions while taking into account the signal variability. We formulate a novel stochastic optimization model and propose a piecewise linearization method for accurately representing the operational decisions' effect on the degradation models in combination with other formulation enhancements. We also extend the chance constraint proposed in Basciftci et al. (2018) to the decision-dependent setting. To evaluate the performances of the maintenance schedules, we develop a decision-dependent simulation framework. This framework enables determining the quality of a solution by simulating signals based on a given schedule. We provide a comprehensive computational study on two illustrative IEEE test cases by comparing the proposed load-dependent approach with load-independent and reliability-based approaches.

We present a computational analysis by optimizing the schedules under different congestion levels and conservativeness amount of the chance constraint. Our analysis demonstrates the superior performances of the load-dependent schedules with reductions in failures and significant cost savings up to 20%. Finally, we provide experiments demonstrating the computational efficiency of the formulation improvements up to 17 times speedup. These results highlight the importance of considering operational decisions in condition-based maintenance scheduling to ensure reliableness and cost effectiveness of the system.

Appendix A Decision-dependent simulation framework

Algorithm 1 Solution Evaluation

Obtain z^*, y^*, x^*, ψ^* solutions from the optimization model (11).
Set $numPaths = 1000$, $numFailures = 0$, $maintCost = 0$, $totalCost = 0$, $FP = firstPhase$.
for all $l \in \{1, \cdots, numPaths\}$ **do**
  for all $i \in G$ **do**
    $Amp_i = \sum_{j=1}^{k^i} S_j^i$, $hasMainted = False$, $hasFailed$ = False, $maintCompPeriod = 0$.
    Generate initial amplitude after maintenance, $newAmp_i$, from the distribution of θ_i.
    for all $t \in \{1, \cdots, H\}$ **do**
      if $z_{i,t}^* == 1$ and $hasFailed ==$ False **then**
        $Amp_i = newAmp_i$, $maintCost \mathrel{+}= C^p$, $hasMainted$= True, $maintCompPeriod = t + Y_p$.
      else if $Amp_i > \Lambda$ **then**
        $Amp_i = newAmp_i$, $numFailures \mathrel{+}= 1$, $maintCost \mathrel{+}= C^c$, $hasMainted = True$, $hasFailed$ = True, $maintCompPeriod = t + Y_c$.
      else if $hasMainted ==$ False or ($t \geq maintCompPeriod + FP$ and $hasMainted == True$) **then**
        Calculate unit degradation d in period t by $d = \sum_{s\in\mathcal{S}}(\alpha_i y_{i,t,s}^* + \beta_i x_{i,t,s}^*)$.
        $Amp_i \mathrel{+}= \mu_i d + \sigma m d$, where m is sampled from $N(0, 1)$.
      end if
      if $t \geq maintCompPeriod$ **then**
        $totalCost \mathrel{+}= \sum_{s\in\mathcal{S}}(V_{i,t,s} y_{i,t,s}^* + F_{i,t,s} x_{i,t,s}^*)$.
      else
        $totalCost \mathrel{+}= \sum_{s\in\mathcal{S}} P_{DC} y_{i,t,s}^*$.
      end if
    end for
  end for
  $totalCost \mathrel{+}= \sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{S}}\sum_{b\in\mathcal{B}} P_{DC}\psi_{b,t,s}^*$.
end for
$totalCost \mathrel{+}= maintCost$.
Divide $numFailures$, $maintCost$, $totalCost$ by $numPaths$ to find the mean results.

Notes on contributors

Beste Basciftci is a postdoctoral fellow in the H. Milton Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology, Atlanta, GA, USA. She completed her

PhD in operations research in the same institution, and received her B.S. and M.S. degrees in industrial engineering and a B.S. degree in computer engineering from Bogazici University, Istanbul, Turkey. Dr. Basciftci's research focuses on data-driven decision making under uncertainty by developing mixed-integer, stochastic programming and distributionally robust optimization approaches to large-scale operations research/management related problems.

Shabbir Ahmed was the Anderson-Interface Chair and Professor in the H. Milton Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology. His research interests were in stochastic and discrete optimization. Dr. Ahmed was a past Chair of the Stochastic Programming Society. Dr. Ahmed's honors include the Farkas Prize from the INFORMS Optimization Society, the INFORMS Computing Society Prize, the National Science Foundation CAREER award, two IBM Faculty Awards, and the INFORMS Dantzig Dissertation award. He was a Senior Member of IEEE and a Fellow of INFORMS.

Nagi Z. Gebraeel is the Georgia Power Early Career Professor and Professor in the H. Milton Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology. His research interests lie at the intersection of industrial predictive analytics and decision optimization models for large scale power generation applications. Dr. Gebraeel serves as an associate director at Georgia Tech's Strategic Energy Institute and the director of the Analytics and Prognostics Systems laboratory at Georgia Tech's Manufacturing Institute. Dr. Gebraeel was the former president of the IIE Quality and Reliability Engineering Division, and is currently a member of INFORMS.

References

Abiri-Jahromi, A., Fotuhi-Firuzabad, M., and Parvania, M. (2012). Optimized midterm preventive maintenance outage scheduling of thermal generating units. *IEEE Transactions on Power Systems*, *27*(3), 1354-1365.

Athay, T., Podmore, R., and Virmani, S. (1979). A practical method for the direct analysis of transient stability. *IEEE Transactions on Power Apparatus and Systems*, *PAS-98*(2), 573-584.

Basciftci, B., Ahmed, S., Gebraeel, N. Z., and Yildirim, M. (2018). Stochastic optimization of maintenance and operations schedules under unexpected failures. *IEEE Transactions on Power Systems*, *33*(6), 6755-6765.

Bian, L., and Gebraeel, N. Z. (2013). Stochastic methodology for prognostics under continuously varying environmental profiles. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *6*(3), 260-270.

Billinton, R., and Mo, R. (2005). Composite system maintenance coordination in a deregulated environment. *IEEE Transactions on Power Systems*, *20*(1), 485-492.

Blumsack, S. (2006). *Network topologies and transmission investment under electric-industry restructuring* (PhD Thesis). Carnegie Mellon University, Pittsburgh, Pennsylvania.

Cain, M. B., O'Neill, R. P., and Castillo, A. (2012). *History of optimal power flow and formulations.* http://www.ferc.gov/industries/electric/indus-act/market-planning/opf-papers/acopf-1-history-formulation-testing.pdf.

Canto, S. P. (2008). Application of benders' decomposition to power plant preventive maintenance scheduling. *European Journal of Operational Research*, *184*(2), 759 - 777.

Conejo, A. J., Garcia-Bertrand, R., and Diaz-Salazar, M. (2005). Generation maintenance scheduling in restructured power systems. *IEEE Transactions on Power Systems*, *20*(2), 984-992.

Doksum, K. A., and Hóyland, A. (1992). Models for variable-stress accelerated life testing experiments based on wiener processes and the inverse gaussian distribution. *Technometrics*, *34*(1), 74-82.

Elwany, A. H., and Gebraeel, N. Z. (2008). Sensor-driven prognostic models for equipment replacement and spare parts inventory. *IIE Transactions*, *40*(7), 629-639.

Froger, A., Gendreau, M., Mendoza, J. E., Pinson, E., and Rousseau, L. (2016). Maintenance scheduling in the electricity industry: A literature review. *European Journal of Operational Research*, *251*(3), 695 - 706.

Gebraeel, N., Lawley, M. A., Li, R., and Ryan, J. K. (2005). Residual-life distributions from component degradation signals: A bayesian approach. *IIE Transactions*, *37*(6), 543-557.

Gebraeel, N., and Pan, J. (2008). Prognostic degradation models for computing and updating residual life distributions in a time-varying environment. *IEEE Trans-*

actions on Reliability, *57*(4), 539-550.

Goel, V., and Grossmann, I. E. (2006). A class of stochastic programs with decision dependent uncertainty. *Mathematical Programming*, *108*(2), 355-394.

Kumar, D., and Klefsjo, B. (1994). Proportional hazards model: a review. *Reliability Engineering and System Safety*, *44*(2), 177 - 188.

Levitin, G., Xing, L., and Dai, Y. (2017). Preventive replacements in real-time standby systems with periodic backups. *IEEE Transactions on Reliability*, *66*(3), 771-782.

Liao, H., and Tian, Z. (2013). A framework for predicting the remaining useful life of a single unit under time-varying operating conditions. *IIE Transactions*, *45*(9), 964-980.

Pourahmadi, F., Fotuhi-Firuzabad, M., and Dehghanian, P. (2017). Application of game theory in reliability-centered maintenance of electric power systems. *IEEE Transactions on Industry Applications*, *53*(2), 936-946.

Si, X. S., Wang, W., Hu, C. H., and Zhou, D. H. (2011). Remaining useful life estimation – a review on the statistical data driven approaches. *European Journal of Operational Research*, *213*(1), 1 - 14.

van Ackooij, W., Danti Lopez, I., Frangioni, A., Lacalandra, F., and Tahanan, M. (2018). *Large-scale unit commitment under uncertainty: an updated literature survey.* University of Pisa.

Vielma, J. P., Ahmed, S., and Nemhauser, G. L. (2010). Mixed-integer models for nonseparable piecewise-linear optimization: Unifying framework and extensions. *Operations Research*, *58*(2), 303-315.

Vielma, J. P., and Nemhauser, G. L. (2008). Modeling disjunctive constraints with a logarithmic number of binary variables and constraints. In *Proceedings of the 13th international conference on integer programming and combinatorial optimization* (pp. 199–213). Berlin, Heidelberg: Springer-Verlag.

Wang, Y., Li, Z., Shahidehpour, M., Wu, L., Guo, C. X., and Zhu, B. (2016). Stochastic co-optimization of midterm and short-term maintenance outage scheduling considering covariates in power systems. *IEEE Transactions on Power Systems*, *31*(6), 4795-4805.

Wilf, H. (1989). *Combinatorial algorithms–an update* (Vol. 55). Society for Industrial and Applied Mathematics.

Wu, L., Shahidehpour, M., and Fu, Y. (2010). Security-constrained generation and transmission outage scheduling with uncertainties. *IEEE Transactions on Power Systems*, *25*(3), 1674-1685.

Wu, L., Shahidehpour, M., and Li, T. (2008). Genco's risk-based maintenance outage scheduling. *IEEE Transactions on Power Systems*, *23*(1), 127-136.

Yildirim, M., Gebraeel, N. Z., and Sun, X. A. (2016a). Sensor-driven condition-based generator maintenance scheduling;part ii: Incorporating operations. *IEEE Transactions on Power Systems*, *31*(6), 4263-4271.

Yildirim, M., Gebraeel, N. Z., and Sun, X. A. (2016b). Sensor-driven condition-based generator maintenance scheduling;part i: Maintenance problem. *IEEE Transactions on Power Systems*, *31*(6), 4253-4262.

Yildirim, M., Gebraeel, N. Z., and Sun, X. A. (2019). Leveraging predictive analytics to control and coordinate operations, asset loading and maintenance. *IEEE Transactions on Power Systems*. doi: 10.1109/TPWRS.2019.2922380

Zheng, Q. P., Wang, J., and Liu, A. L. (2015). Stochastic optimization for unit commitment; a review. *IEEE Transactions on Power Systems*, *30*(4), 1913-1924.