Condition-based Maintenance for Wind Farms using a Distributionally Robust Chance Constrained Program

Heraldo Rozas, Weijun Xie, Nagi Gebraeel

Abstract—**Operations and Maintenance (O&M) expenses account for up to 30% of the operating costs of wind farms. Condition-based maintenance (CBM) strategies, which incorporate predictive analytics into maintenance optimization, have been proven to be effective in reducing O&M costs in wind farms. Existing predictive CBM strategies for wind farms rely on the assumption that predictive analytics can accurately estimate the remaining lifetime distribution (RLD) of wind turbines, allowing for the direct implementation of stochastic programming or threshold-based policies. However, estimated RLDs can be inaccurate due to noisy sensors or limited training data. To address this issue, this paper develops a CBM strategy for wind farms that uses a Distributionally Robust Chance Constrained (DRCC) optimization model. Our formulation acknowledges that estimated distributions may be inaccurate and so seeks solutions that are robust against distribution perturbations within a Wasserstein ambiguity set. We show that the proposed DRCC optimization problem can be exactly reformulated as an integer linear program. We derive methods to strengthen the Big-M values of this reformulation, thereby enabling the DRCC model to be efficiently solved by off-the-shelf optimization software. The proposed strategy is validated through computational studies using real-world and synthetic degradation data, outperforming stochastic programming and robust optimization benchmark models.**

Index Terms—**Wind Turbines, Distributionally Robust Chance Constrained Optimization, Condition-based Maintenance, Condition Monitoring, Prognostics.**

Nomenclature

Abbreviations

CM	Condition monitoring.
CBM	Condition-based maintenance.
DR	Distributionally robust.
DRCC	Distributionally robust chance constrained.
DRO	Distributionally robust optimization.
IP	Integer programming.
MILP	Mixed integer linear programming.
O&M	Operations and Maintenance.
RLD	Remaining lifetime distribution.
SAA	Sample average approximation.
WF	Wind farm.
WT	Wind turbine.

Sets

$\mathcal{J}$	Set of wind turbines.
$\mathcal{T}$	Set of maintenance epochs.
$\mathcal{N}$	Set of samples.
$\mathcal{P}^W(\delta)$	Wasserstein ambiguity set with radius δ.
$\mathcal{X}$	Feasible set of decision variables.
$\mathcal{D}$	Feasible set of decision variables determined by the deterministic constraints.

Decision variables

x_{jt}	$x_{jt} = 1$ if wind turbine j undergoes maintenance at time t.
z_t	$z_t = 1$ if maintenance crew is deployed in the wind farm at time t.
$\boldsymbol{X}$	Vector of all decision variables.

Random variables and distributions

$\tilde{\omega}_j$	Random variable representing the remaining lifetime of wind turbine j.
$\tilde{\Omega}$	Random vector representing the remaining lifetimes of all wind turbines under study.
RLD_j	Remaining lifetime distribution of wind turbine j.
$\widehat{RLD}_j$	Estimated remaining lifetime distribution of wind turbine j.

Parameters

J	Total number of wind turbines.
N	Total number of samples.
c^{pr}	Preventive maintenance cost.
c^{co}	Corrective maintenance cost.
c^v	Deployment cost of the maintenance crew.
p_t	Expected revenue of wind turbine at time t.
K_t	Capacity of the maintenance crew at time t.
T^{max}	Length of the planning horizon.
Δ^{upd}	Length of the freeze period.
ρ_j	Threshold on the maximum unavailability of wind turbine j.
λ_t	Threshold on the maximum number of unavailable wind turbines at time t.
γ	Threshold on the maximum number of wind turbine failures.
ϵ	Confidence level of the first chance constraint.
α	Confidence level of the second chance constraint.
β	Confidence level of the third chance constraint.
δ	Radius of the Wasserstein ambiguity set.
ω_j^i	Remaining lifetime sample i of wind turbine j.
Ω_i	Remaining lifetime vector sample i of all wind turbines under study.
σ_j	Empirical standard deviation of $\{\omega_j^i\}_{i\in\mathcal{N}}$.

The first and third authors were supported by NASA under Grant 80NSSC19K1052 as part of the NASA Space Technology Research Institute (STRI) Habitats Optimized for Missions of Exploration (HOME) SmartHab Q3 Project. The second author was supported by NSF CMMI Grant # 2246414.

The authors are with the School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA (e-mails: heraldo.rozas@gatech.edu, wxie@gatech.edu, nagi.gebraeel@isye.gatech.edu)

I. Introduction

Wind power generation has been growing steadily in the United States, constituting 30% of total capacity additions over the last decade [1]. Reliable power generation from wind farms (WFs) relies heavily on the Operations and Maintenance (O&M) strategies. It is estimated that O&M activities account for up to 30% of the running costs of WF operations [2], [3]. Therefore, enhancing maintenance and repair operations can significantly impact the bottom-line profits of WFs. This need has driven extensive research into optimization strategies for WF O&M, as discussed in several survey papers [2]–[4].

One of the key challenges in optimizing WF O&M is the uncertainty that exists around the lifetime distribution of wind turbines (WTs). This uncertainty poses a significant challenge for stakeholders who are responsible for managing the O&M of WFs, as it can impact the overall efficiency and profitability of the operation. The impact of uncertainty on the remaining lifetime of WTs and their maintenance schedules is a critical factor that is frequently overlooked in the planning and operational phases of WFs. While some studies may assume that the remaining lifetime of WTs has been accurately estimated [5] or that maintenance windows are predetermined [6]–[9], this approach can lead to underestimation of the true scope of uncertainty involved. In reality, a significant body of research acknowledges the importance of considering the uncertainty in WT lifetimes due to its substantial effect on maintenance costs and the overall reliability of WFs.

One of the common approaches to dealing with lifetime uncertainty when optimizing maintenance schedules for WTs is by using reliability models. These models make use of historical data on failure times to fit standard reliability distributions, such as the Weibull distribution, which help to describe the uncertainty surrounding the failure times of WTs [10], [11]. Fitted distributions can then be integrated into the optimization problem to calculate appropriate maintenance schedules. These optimization problems seek to minimize maintenance costs. They are often framed as reliability/age threshold policies, wherein maintenance actions are performed when the WT reliability/age surpasses a predetermined threshold [12]–[20]. Previous reliability-driven formulations have explored various maintenance strategies, including preventive maintenance (conducted before failure), corrective maintenance (conducted after failure), and opportunistic maintenance [12], [17]–[19]. Opportunistic maintenance involves grouping maintenance tasks for WT components based on economic convenience, such as aligning with already scheduled maintenance tasks or favorable weather conditions [21]. Additional operational aspects of reliability-based WT maintenance optimization include spare parts management [18], [22] and maintenance crew routing [23].

In contrast, condition-based maintenance (CBM) leverages condition monitoring (CM) data to optimize maintenance schedules, adapting maintenance decisions to the observed degradation of WTs [24]. CBM has received significant attention for its potential to reduce the operational costs of WFs and has been investigated using two different approaches: the state-based approach and the predictive analytics-based approach. The state-based approach assumes that the current degradation state of the WT can be estimated/observed using CM technology. The collected information can be used to compute efficient maintenance schedules that better reflect the state of the WT relative to the reliability-based approach discussed earlier. For example, in [25], [26], the authors developed a partially observed Markov decision process (POMDP) to compute WT maintenance schedules. Their model considers seasonal weather variations and assumes that the exact degradation states of the WTs can be observed by paying an observation cost. In [27], the authors modeled the degradation of WTs using a continuous-time Markov chain with discrete states, which can be estimated by either visual or remote inspections. The inspection policy was optimized using Monte Carlo simulation. State-based approaches, like the ones discussed, utilized sensor data to estimate the current WT degradation state. However, they did not predict its future evolution.

On the other hand, predictive analytics approaches rely on prognostic models that use sensor data to estimate the remaining lifetime distributions (RLDs) of WTs. They focus on predicting the future degradation state of the WT. Estimated RLDs generally exhibit lower levels of uncertainty compared to reliability models, primarily because RLD predictions are based on actual degradation signals gathered from operational WTs. These observed signals provide a more concrete basis for estimation, thus reducing the uncertainty compared to estimates based on historical lifetime data. By incorporating these estimated RLDs into optimization modules, it is possible to devise more efficient maintenance schedules for WTs.

Advances in predictive analytic models for WTs have been notable during the last decades, as surveyed in the following review papers [28]–[31]. However, few works have tackled the integration of predictive analytics and optimization within a single framework that aims to compute the maintenance schedule of WTs. These frameworks can be classified into two categories: threshold-based policies and stochastic programming models. In threshold-based policies such as those found in [32]–[37] aim to construct maintenance policies that are based on a set of decision parameters, such as intervals between repairs or signal thresholds. They evaluate the expected performance of the policy for different decision parameters using numerical methods built with the RLDs estimated by prognostic models. The optimal maintenance policy is determined by selecting the parameters that provide the best expected performance (e.g., minimum maintenance cost or maximum revenue). Stochastic programming models such as [38]–[44] focus on formulating stochastic mixed integer linear programs to minimize the expected maintenance cost or maximize the expected total revenue while satisfying operational constraints. The stochasticity of the optimization model is driven by the RLDs that are estimated using a prognostic model. Some authors [40], [41], [43] include chance constraints to restrict the probability of undesired events, such as a large number of simultaneous unavailable generators or WTs.

The direct integration of estimated RLDs into decision frameworks, as demonstrated by the previous works, relies on the underlying assumption that the estimated RLDs accurately represent the "true" RLDs. In other words, these works assume that the prognostic model generates accurate predictions. Unfortunately, this assumption often does not hold true in practice due to multiple factors, including noisy sensors or sparsity of (historical) training data. In these scenarios, there is no guarantee of how well these models will perform [24].

This paper addresses this problem by proposing and developing a CBM model for WFs that is based on a Distributionally Robust Chance Constrained (DRCC) formulation. Our model focuses on the uncertainty of estimating the remaining lifetimes of WTs. We utilize a prognostic model that utilizes real-time sensor measurements to update the RLDs in real-time. Consequently, we implement our optimization model in a rolling horizon fashion to adapt its decisions to the latest degradation states of the individual WTs. Our proposed DRCC formulation acknowledges that the estimated RLDs can be inaccurate. It, therefore, seeks optimal robust solutions against perturbations of the RLD within a Wasserstein ambiguity set. Our formulation aims to optimize maintenance schedules in the presence of inaccurate prognostic results. Our objective function aims to minimize expected maintenance costs, accounting for the trade-off between early and late repairs. Our model encourages opportunistic repairs by incorporating the deployment cost of the maintenance crew into our objective function. This serves as an incentive to group maintenance activities whenever it is economically convenient. Our formulation also includes operational contract requirements by incorporating Distributionally Robust (DR) chance constraints. Specifically, we formulate and implement three different DR chance constraints. The first constraint aims to limit the unavailability of WTs. The second is designed to ensure that the power generation commitments are satisfied. The third constraint restricts the number of corrective repairs (resulting from unexpected failures). We show that the proposed DRCC optimization problem can be exactly reformulated as an integer program that can be solved efficiently using off-the-shelf solvers. In summary, the contributions of this paper are threefold:

- We develop a CBM optimization model for WFs that is based on a DRCC formulation. Our model aims to minimize expected maintenance costs while incorporating operational and reliability requirements. We utilize a contemporary prognostic model that utilizes real-time degradation-based sensor data to predict and continuously update RLDs of WTs. Unlike existing models, we account for potential estimation errors and uncertainty in the predicted RLDs of individual WTs due to noise and/or sparsity of the training data. We propose a CBM model that relies on a DR formulation where the predicted RLDs are utilized to construct a Wasserstein ambiguity set that models potential perturbations of the estimated distribution. To the best of our knowledge, this type of formulation is the first of its kind in WF maintenance literature.
- We formulate three data-driven DR chance constraints that aim to limit the probability associated with extended unavailability of WTs, not meeting power demand, and having a large number of unexpected failures. The simultaneous implementation of these three DR chance constraints is unique to this work and can be helpful for operators to verify that the computed maintenance schedules satisfy reliability and operational contract requirements. Such verification cannot be achieved if we solely focus on minimizing expected maintenance costs, which is often the case in the existing literature.
- We derive an exact integer programming (IP) reformulation for the proposed DRCC maintenance optimization problem by exploiting specific structural properties of our problem. We provide closed-form expressions to compute the parameters of the resulting IP reformulation. Furthermore, we develop a methodology to strengthen the Big-M parameters of the reformulation, enabling efficient solving of the model by off-the-shelf optimization software. The performance of the proposed DRCC formulation is evaluated through simulation studies built with real-world and synthetic degradation data. The proposed DR formulation shows results superior to the benchmark models based on stochastic programming and robust optimization, especially when dealing with inaccurate prognostic results.

The remainder of this paper is organized as follows. Section II presents the problem setting. Section III details the proposed DRCC formulation and the corresponding IP reformulation. Section IV discusses the results obtained in the computational studies. Finally, Section V provides the final remarks and future research directions of this work.

II. Problem Setting

We consider a fleet of WTs indexed by $j \in \mathcal{J} = \{1, ..., J\}$. We assume that these WTs are monitored by sensors and that the collected CM data can be used to predict the remaining operational life of the WT. The remaining lifetime of the $j-$th WT is defined as a non-negative random variable, denoted as $\tilde{\omega}_j$. We denote $\tilde{\Omega} = [\tilde{\omega}_1, ..., \tilde{\omega}_J]$ as a random vector representing the remaining lifetimes for all WTs. Note that WTs are assumed to be independent. We consider a setting where a prognostic model is used to analyze (streaming) CM data to predict (and update) the RLD of the WT. The resulting distribution is denoted as $\widehat{RLD}_j$ and corresponds to a data-driven estimate of the ground truth distribution, RLD_j, which governs the uncertainty of the component remaining lifetime $\tilde{\omega}_j$.

As we noted earlier, only a handful of works have truly integrated RLD predictions within their maintenance optimization models, such as [32], [34], [35], [39]–[44]. However, they all assumed that the predicted RLDs are accurate. In practice, this is not entirely true. Prognostic models leverage noisy data to predict RLDs. Additionally, many prognostic models are estimated using historical data that is often sparse and fragmented. Consequently, the prognostic model being used and the resulting remaining life estimates are likely going to be inaccurate. Due to this uncertainty, optimization models that use methods such as stochastic programming and simulation-based approaches cannot guarantee good performance in these settings.

To address these challenges, we propose a DRCC formulation that seeks optimal solutions that are robust against distribution perturbations within a Wasserstein ambiguity set. Another attractive feature of our model is that it is not bound to a particular prognostic modeling methodology. In other words, any prognostic modeling approach that predicts RLDs can be easily integrated with our CBM optimization model. Our formulation assumes that (1) each WT needs a single repair within the planning horizon; (2) if a WT has been repaired within the planning horizon, it cannot fail within the remaining periods of the same planning horizon; (3) power demand, power generation, and energy price are assumed to be known; and (4) a WT can be inactive due to two reasons only, failure or ongoing maintenance.

The implementation of the proposed maintenance strategy with its different building blocks is summarized in Figure 1. Sensor data collected from WTs are processed by a prognostic model to estimate $\widehat{RLD}_j,\ j \in \mathcal{J}$. These estimated RLDs are then integrated into the proposed DRCC formulation, which is presented in the next section, to compute the maintenance of WTs.

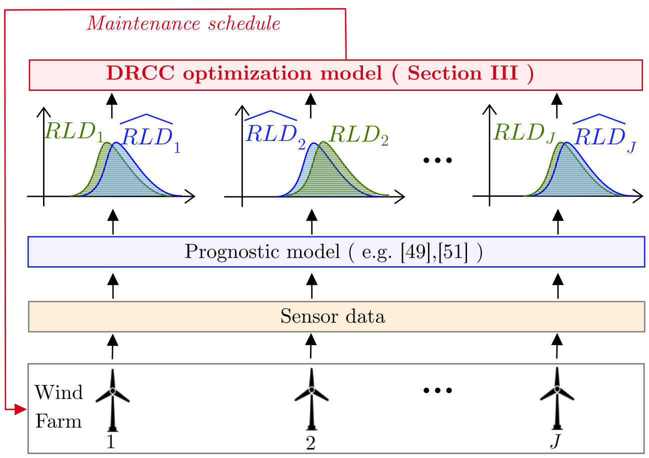

Fig. 1: Proposed maintenance strategy.

III. A Distributionally Robust Chance Constrained Formulation

The first step toward building our DRCC formulation is to define the type of distribution perturbation allowed within the optimization framework. This is modeled by the ambiguity set presented in the following section.

A. Ambiguity Set

The uncertainty of the CBM optimization model lies in the remaining lifetime of the J WTs, which can be represented as a vector of non-negative independent random variables denoted as $\tilde{\Omega} = [\tilde{\omega}_1, ..., \tilde{\omega}_J]$. The distribution of $\tilde{\omega}_j$ is unknown, but it can be estimated using a prognostic model, resulting in $\widehat{RLD}_j$. By assuming the independence of $\tilde{\omega}_j$ for all $j \in \mathcal{J}$, we can estimate the joint distribution of $\tilde{\Omega}$ and denote it as $\widehat{RLD}$.

Next, we construct an empirical distribution of $\tilde{\Omega}$ by generating N independent samples of $\widehat{RLD}$. These samples are indexed by $i \in \mathcal{N} = 1, ..., N$, and the i-th sample is denoted as $\Omega_i = [\omega_1^i, ..., \omega_J^i]$, where each ω_j^i is drawn from $\widehat{RLD}_j$. Since the age of WTs can vary significantly, there might be substantial differences in the range of random variables representing the remaining lifetime of each WT. To standardize this variability, we define a normalized sample of remaining lifetimes as $\Omega_i^{'} = [{\omega'}_1^i, ..., {\omega'}_J^i]$, where ${\omega'}_j^i = \omega_j^i/\sigma_j$, and σ_j is the empirical standard deviation of $\{\omega_j^i\}_{i\in\mathcal{N}}$. The empirical distribution based on the normalized set of samples $\{\Omega_i^{'}\}_{i\in\mathcal{N}}$ is given by

$$\hat{\mathbb{P}}_N(\tilde{\Omega}^{'}) = \frac{1}{N}\sum_{i\in\mathcal{N}} \mathbb{I}(\tilde{\Omega}^{'} = \Omega_i^{'}),$$

where $\mathbb{I}(x)$ is the indicator function. Our approach uses the $\infty-$Wasserstein ambiguity set [45] defined as:

$$\mathcal{P}^W(\delta) = \left\{\mathbb{P} \in \mathcal{P}(\Xi) : W_\infty\left(\mathbb{P}, \hat{\mathbb{P}}_N\right) \le \delta\right\}, \qquad (1)$$

where the $\infty-$Wasserstein distance is given by:

$$W_\infty(\mathbb{P}_1, \mathbb{P}_2) = \inf_{\mathbb{Q}}\left\{ \text{ess.sup}\|\tilde{\Omega}_1^{'} - \tilde{\Omega}_2^{'}\| : \begin{array}{l} \mathbb{Q} \text{ is a joint distribution of } \tilde{\Omega}_1^{'} \text{ and } \tilde{\Omega}_2^{'} \\ \text{with marginals } \mathbb{P}_1 \text{ and } \mathbb{P}_2\text{, respectively} \end{array}\right\}.$$

$\mathcal{P}(\Xi)$ represents the set of all distributions with support $\Xi = \text{supp}(\tilde{\Omega}^{'})$. Therefore, $\mathcal{P}^W(\delta)$ represents the set of distributions within Ξ whose $\infty-$Wasserstein distance from the empirical distribution $\hat{\mathbb{P}}_N$ is less than or equal to $\delta \ge 0$. This set can be visualized as a Wasserstein ball centered at $\hat{\mathbb{P}}_N$ with a radius of δ. The hyper-parameter $\delta \ge 0$ controls the size of the ambiguity set and determines the level of confidence in the prognostic model used to predict the RLDs. A smaller δ indicates a higher confidence level in the prognostic results, while a larger δ implies a lower confidence level. The adoption of a Wasserstein ambiguity set is driven by two practical reasons: i) it allows us to restrict the shape of the distribution perturbations for general empirical baseline distributions, and ii) it enables us to derive tractable reformulations that can be solved directly with off-the-shelf optimization software. These two reasons are common criteria for selecting ambiguity sets in distributionally robust optimization (DRO) applications [46]. For the derivations of the next section, we will use the maximum norm that is defined and denoted as follows: $\|v\| = \max_{i\in[1,...,d]}|v_i|,\ v \in \mathbb{R}^d$.

B. Decision Variables and Objective Function

This section discusses the decision variables and the objective function of our formulation. Let $\mathcal{T} = \{1, 2, ..., T^{max}\}$ be the set of maintenance epochs and $\mathcal{J}$ be the set of WTs. We introduce the binary decision variable x_{jt}, where $x_{jt} = 1$ if WT $j \in \mathcal{J}$ is scheduled to be repaired at time $t \in \mathcal{T}$. We also define a binary decision variable z_t, $t \in \mathcal{T}$, where $z_t = 1$ if the maintenance crew visits the WF at time t. This decision variable is used to capture the deployment or setup cost of the maintenance crew. This cost is similar to the cost used by [39] to encourage opportunistic repairs. Opportunistic repairs are very attractive when the cost of allocating and deploying maintenance resources is high.

Our objective is to minimize the total maintenance costs. To this end, we propose the cost function expressed in (2), where

$\boldsymbol{X}$ is a vector that contains all decision variables. The cost function is divided into three parts. The first part represents a preventive maintenance cost incurred when WT j is repaired at time t prior to failure (i.e., $\tilde{\omega}_j > t$). This cost consists of a fixed repair cost c^{pr} plus a variable opportunity cost $\sum_{k=t}^{\lfloor\tilde{\omega}_j\rfloor-1} p_k$ that is used to account for the lost production due to the early repair. Note that this is a sunk cost that becomes higher when WTs are maintained prematurely, i.e., their remaining lifetimes are significant. The second component is a corrective maintenance cost incurred at time t when WT j is repaired after failure, i.e., $\tilde{\omega}_j \leq t$. This cost also consists of a fixed repair cost c^{co} (where $c^{co} > c^{pr}$) plus the variable opportunity cost $\sum_{k=\lfloor\tilde{\omega}_j\rfloor}^{t-1} p_k$, which again captures the lost production. The third cost component is the setup cost. It is used to capture the deployment cost, c^v, of the maintenance crew. As mentioned earlier, this cost is critical to encouraging opportunistic maintenance by grouping the repair operations of multiple WTs at the same maintenance event. This cost attempts to balance the previous two costs components with respect to the cost of deploying the crew to a given WF [39].

It is worth mentioning that our model does not schedule maintenance upon the failure of individual WTs. Instead, it seeks to compute a maintenance schedule that balances between early and late repairs for the entire fleet of WTs, taking into account penalties associated with generation loss and setup costs. Consequently, the model may postpone the repair of a failed WT to align it with the repair of other degraded WTs when such alignment is more economically convenient. The model incorporates specific DR chance constraints to ensure that the unavailability of WTs (i.e., postponing the repair of failed WTs) does not compromise the operational requirements of the WF–these constraints are discussed later in Section III-D.

$$
\begin{aligned}
f(\boldsymbol{X},\tilde{\Omega}) = & \overbrace{\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\left(\left(\sum_{k=t}^{\lfloor\tilde{\omega}_j\rfloor-1} p_k + c^{pr}\right)\mathbb{I}\{\tilde{\omega}_j > t\}\right)\cdot x_{jt}+}^{\text{(i) Preventive Maintenance Cost}} \\
& \overbrace{\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\left(\left(\sum_{k=\lfloor\tilde{\omega}_j\rfloor}^{t-1} p_k + c^{co}\right)\mathbb{I}\{\tilde{\omega}_j \leq t\}\right)\cdot x_{jt}+}^{\text{(ii) Corrective Maintenance Cost}} \\
& + \overbrace{\sum_{t\in\mathcal{T}} c^v\cdot z_t}^{\text{(iii) Setup Cost}}
\end{aligned}
\tag{2}
$$

Note that $f(\boldsymbol{X},\tilde{\Omega})$ can be rewritten as follows:

$$
f(\boldsymbol{X},\tilde{\Omega}) = \sum_{t\in\mathcal{T}} c^v\cdot z_t + \sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\alpha(\tilde{\omega}_j,t)\cdot x_{jt},
$$

where $\alpha(\tilde{\omega}_j,t) = \left(\sum_{k=t}^{\lfloor\tilde{\omega}_j\rfloor-1} p_k + c^{pr}\right)\mathbb{I}\{\tilde{\omega}_j > t\} + \left(\sum_{k=\lfloor\tilde{\omega}_j\rfloor}^{t-1} p_k + c^{co}\right)\mathbb{I}\{\tilde{\omega}_j \leq t\}$.

Note that the cost function $f(\boldsymbol{X},\tilde{\Omega})$ is stochastic due to its dependence on $\tilde{\Omega}$. Our goal is to solve the following DRO problem:

$$
\min_{\boldsymbol{X}\in\mathcal{X}}\left\{\sup_{\mathbb{P}\in\mathcal{P}^W(\delta)}\mathbb{E}^{\mathbb{P}}[f(\boldsymbol{X},\tilde{\Omega})]\right\}, \tag{3}
$$

where $\mathcal{X}\subseteq\mathbb{R}^{|\boldsymbol{X}|}$ corresponds to the feasible region of $\boldsymbol{X}$ determined by the operational constraints explained later.

We now focus on deriving a tractable expression for $\sup_{\mathbb{P}\in\mathcal{P}^W(\delta)}\mathbb{E}^{\mathbb{P}}[f(\boldsymbol{X},\tilde{\Omega})]$. To this end, we notice that (3) can be simplified using the fact that $\sum_{t\in\mathcal{T}} c^v\cdot z_t$ is a deterministic cost. This results in the following equation:

$$
\begin{aligned}
\sup_{\mathbb{P}\in\mathcal{P}^W(\delta)}\mathbb{E}^{\mathbb{P}}[f(\boldsymbol{X},\tilde{\Omega})] = & \sum_{t\in\mathcal{T}} c^v\cdot z_t + \\
& \sup_{\mathbb{P}\in\mathcal{P}^W(\delta)}\mathbb{E}^{\mathbb{P}}\left[\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\alpha(\tilde{\omega}_j,t)\cdot x_{jt}\right].
\end{aligned}
\tag{4}
$$

Before detailing our derivations, we present two results in Preposition 1 and 2 adapted from [47] and [45], respectively. These results are stated in terms of the notation used in our model.

Proposition 1. *(Adapted from Proposition 3 in [47]) Consider a real-valued function $r(\boldsymbol{X},\tilde{\Omega}^{'})$ dependent on the vector of decision variables $\boldsymbol{X}$ and the random vector $\tilde{\Omega}^{'}$. Then,*

$$
\min_{\boldsymbol{X}\in\mathcal{X}}\left\{\sup_{\mathbb{P}\in\mathcal{P}^W(\delta)}\mathbb{E}^{\mathbb{P}}[r(\boldsymbol{X},\tilde{\Omega}^{'})]\right\} = \min_{\boldsymbol{X}\in\mathcal{X}}\left\{\frac{1}{N}\sum_{i\in\mathcal{N}}\sup_{\|\Omega^{'}-\Omega_i^{'}\|\leq\delta} r(\boldsymbol{X},\Omega^{'})\right\}.
$$

Proposition 2. *(Adapted from Proposition 3 in [45]) Consider a real-valued function $r(\boldsymbol{X},\tilde{\Omega}^{'})$ dependent on the vector of decision variables $\boldsymbol{X}$ and the random vector $\tilde{\Omega}^{'}$. Then,*

$$
\begin{aligned}
\inf_{\mathbb{P}\in\mathcal{P}^W(\delta)}\mathbb{P}\left\{\tilde{\Omega}^{'} : r(\boldsymbol{X},\tilde{\Omega}^{'})\leq 0\right\} = & \\
& \frac{1}{N}\sum_{i\in\mathcal{N}}\inf_{\|\Omega^{'}-\Omega_i^{'}\|\leq\delta}\mathbb{I}\left(r(\boldsymbol{X},\Omega^{'})\leq 0\right).
\end{aligned}
$$

Now, we proceed to derive a tractable expression of the supremum on the right-hand side of (4). This result is presented by Proposition 3.

Proposition 3. *If $\sum_{t\in\mathcal{T}} x_{jt} = 1$ for any $j\in\mathcal{J}$, then*

$$
\sup_{\mathbb{P}\in\mathcal{P}^W(\delta)}\mathbb{E}^{\mathbb{P}}\left[\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\alpha(\tilde{\omega}_j,t)\cdot x_{jt}\right] = \sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\psi_{jt}\cdot x_{jt},
$$

where $\{\psi_{jt}\}_{j\in\mathcal{J},\, t\in\mathcal{T}}$ are constant values that can be pre-computed from the set of samples $\{\Omega_i\}_{i\in\mathcal{N}}$.

Proof. Using Proposition 1, we obtain the expression in (5a). Since we are using the maximum norm, we can compute the supremum of each WT individually as shown in equation (5b). Since we assume that a WT undergoes only one maintenance event within the planning horizon, then for $\sum_{t\in\mathcal{T}} x_{jt} = 1$, we have only one $x_{jt} = 1$. Thus, the supremum evaluated at time t can be expressed by equation (5c). By rearranging the order of the summation, we arrive at expression (5d).

$$\sup_{\mathbb{P}\in\mathcal{P}^W(\delta)} \mathbb{E}^{\mathbb{P}}\left[\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\alpha(\tilde{\omega}_j,t)\cdot x_{jt}\right]$$

$$=\frac{1}{N}\sum_{i\in\mathcal{N}}\sup_{\Omega':\|\Omega'-\Omega'_i\|\leq\delta}\left(\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\alpha(\omega_j,t)\cdot x_{jt}\right) \tag{5a}$$

$$=\frac{1}{N}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{J}}\sup_{\omega_j:|\omega_j-\omega_j^i|\leq\delta\cdot\sigma_j}\left(\sum_{t\in\mathcal{T}}\alpha(\omega_j,t)\cdot x_{jt}\right) \tag{5b}$$

$$=\frac{1}{N}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\left(\sup_{\omega_j:|\omega_j-\omega_j^i|\leq\delta\cdot\sigma_j}\alpha(\omega_j,t)\right)\cdot x_{jt} \tag{5c}$$

$$=\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\left(\frac{1}{N}\sum_{i\in\mathcal{N}}\sup_{\omega_j:|\omega_j-\omega_j^i|\leq\delta\cdot\sigma_j}\alpha(\omega_j,t)\right)\cdot x_{jt} \tag{5d}$$

Now define ψ_{ijt} as follows:

$$\psi_{ijt}=\sup_{\omega_j:|\omega_j-\omega_j^i|\leq\delta\cdot\sigma_j}(\alpha(\omega_j,t))=$$

$$\max\left\{\begin{array}{ll}\max\sum_{k=t}^{\lfloor\omega_j\rfloor-1}p_k+c^{pr} & \max\sum_{k=\lfloor\omega_j\rfloor}^{t-1}p_k+c^{co}\\ \text{st. } \omega_j>t & , \quad \text{st. } 0\leq\omega_j\leq t\\ \quad |\omega_j-\omega_j^i|\leq\delta\cdot\sigma_j & \quad |\omega_j-\omega_j^i|\leq\delta\cdot\sigma_j\end{array}\right\}. \tag{6}$$

As $p_k\geq 0$, (6) can be further simplified as follows:

$$\begin{aligned}\psi_{ijt}=\max\Bigg\{&\left(\sum_{k=t}^{\Theta_{ij}^{max}}p_k+c^{pr}\right)\cdot\mathbb{I}\left(\omega_j^i+\delta\cdot\sigma_j>t\right),\\ &\left(\sum_{k=\Theta_{ij}^{min}}^{t-1}p_k+c^{co}\right)\cdot\mathbb{I}\left(\omega_j^i+\delta\cdot\sigma_j\leq t\right)\Bigg\},\end{aligned} \tag{7}$$

where $\Theta_{ij}^{max}=\lfloor\min\{\omega_j^i+\delta\cdot\sigma_j-1,T^{max}\}\rfloor$ and $\Theta_{ij}^{min}=\lfloor\max\{\omega_j^i-\delta\cdot\sigma_j,1\}\rfloor$. The proof is completed by setting $\psi_{jt}=\frac{1}{N}\sum_{i\in\mathcal{N}}\psi_{ijt}$, which is a constant that can be pre-computed from the set of samples $\{\Omega_i\}_{i\in\mathcal{N}}$.

□

Remark 1. *Notice that ψ_{ijt} defined in (7) depends on the parameter p_t, which represents the expected revenue generated by a WT at time t and is assumed to be known. In practice, p_t is a random variable due to its dependence on the power generation and energy price. The parameter p_t could be included as a random variable in the proposed DRO maintenance strategy. Doing that would require maximizing ψ_{ijt} with respect to p_k, which is not difficult because ψ_{ijt} is increasing with respect to p_k. However, analyzing the uncertainty of p_t is out of the scope of this paper and will be investigated in future works.*

As our formulation assumes just one repair for each WT within the planning, the assumption of Proposition 3 is satisfied. Therefore, we can invoke Proposition 3 to show that our optimization problem expressed in (3) reduces to the expression below, which corresponds to a linear objective function minimized over $\mathcal{X}$.

$$\min_{\boldsymbol{X}\in\mathcal{X}}\left\{\sum_{t\in\mathcal{T}}c^v\cdot z_t+\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\psi_{jt}\cdot x_{jt}\right\}. \tag{8}$$

The subsequent subsections present the constraints used to describe feasible region $\mathcal{X}$. We will discuss a set of deterministic constraints concerning the number of repairs in the planning horizon and the available capacity of the crew. We will also introduce three different DR chance constraints: the first concerns the unavailability of the WTs, the second concerns demand fulfillment, and the third concerns the number of corrective repairs (resulting from unexpected failures).

C. Deterministic Constraints

We introduce two deterministic constraints modeling some conditions imposed on the maintenance schedule. The feasible region created by these constraints is denoted as $\mathcal{D}$.

Constraint (9) states that any WT needs exactly one repair within the planning horizon.

$$\sum_{t\in\mathcal{T}}x_{jt}=1,\quad j\in\mathcal{J}. \tag{9}$$

Constraint (10) models the deployment of the maintenance crew to conduct maintenance activities in the WF. Specifically, two conditions are captured by this constraint: (1) maintenance tasks can be scheduled at time t only if the maintenance crew visits the WF at time t, and (2) the total number of repairs scheduled at time t cannot exceed the maximum working capacity of the maintenance crew.

$$\sum_{j\in\mathcal{J}}x_{jt}\leq z_t\cdot K_t,\ t\in\mathcal{T}. \tag{10}$$

D. Distributionally Robust Chance Constraints

In addition to deterministic constraints, our formulation includes DR chance constraints to model operational contract requirements associated with reliability indicators and meeting power demand. Upon initial examination,the resulting DRCC sets may seem to be intractable because their construction requires computing $\inf_{\mathbb{P}\in\mathcal{P}^W(\delta)}\mathbb{P}(\cdot)$. We will show that the proposed DRCC sets indeed admit IP reformulations.

1) Limiting the Unavailability of Wind Turbines: We restrict the total number of unavailable time units within the planning horizon for every WT. This is enforced by the following DRCC set:

$$\mathcal{Z}_1=\left\{\boldsymbol{X}\in\mathcal{D}:\inf_{\mathbb{P}\in\mathcal{P}^W(\delta)}\mathbb{P}\left\{\tilde{\Omega}:\ \sum_{t\in\mathcal{T}}[t-\tilde{\omega}_j]_+\cdot x_{jt}\leq\rho_j\right\}\geq 1-\epsilon_j,\quad j\in\mathcal{J}\right\}.$$

Proposition 4. *The set $\mathcal{Z}_1$ admits an IP representation with following structure:*

$$\mathcal{Z}_1=\{\boldsymbol{X}\in\mathcal{D}:x_{jt}\leq u_{jt},\quad j\in\mathcal{J},\ t\in\mathcal{T}\}, \tag{11}$$

where $\{u_{jt}\}_{j\in\mathcal{J},\,t\in\mathcal{T}}$ are binary constant parameters that can be pre-computed using $\{\Omega_i\}_{i\in\mathcal{N}}$.

Proof. Using Proposition 2, we obtain that for a given $j \in \mathcal{J}$:

$$\mathcal{Z}_1=\Bigg\{\boldsymbol{X} \in \mathcal{D}\colon \frac{1}{N}\sum_{i\in\mathcal{N}} \inf_{\omega_j:|\omega_j-\omega_j^i|\le\delta\cdot\sigma_j} \mathbb{I}\left[\sum_{t\in\mathcal{T}}[t-\omega_j]_+ x_{jt} \le \rho_j\right] \ge 1-\epsilon_j, \quad j\in\mathcal{J}\Bigg\}.$$

Since we assume that a WT undergoes only one maintenance event within the planning horizon, i.e., $\sum_{t\in\mathcal{T}} x_{jt} = 1$, we have only one $x_{jt} = 1$. Let $\boldsymbol{X} \in \mathcal{D}$ such that $x_{jt} = 1$. It follows that $\boldsymbol{X} \in \mathcal{Z}_1$ if and only if (12) is satisfied.

$$\frac{1}{N}\sum_{i\in\mathcal{N}} \inf_{\omega_j:|\omega_j-\omega_j^i|\le\delta\cdot\sigma_j} \mathbb{I}\left\{[t-\omega_j]_+ \le \rho_j\right\} \ge 1-\epsilon_j. \tag{12}$$

Next, we can define u_{jt} as follows:

$$u_{jt} = \mathbb{I}\left[\frac{1}{N}\sum_{i\in\mathcal{N}} \inf_{\omega_j:|\omega_j-\omega_j^i|\le\delta\cdot\sigma_j} \mathbb{I}\left\{[t-\omega_j]_+ \le \rho_j\right\} \ge 1-\epsilon_j\right]$$
$$= \mathbb{I}\left[\frac{1}{N}\sum_{i\in\mathcal{N}} \mathbb{I}\left\{[t-\max(\omega_j^i-\delta\cdot\sigma_j,0)]_+ \le \rho_j\right\} \ge 1-\epsilon_j\right],$$

Note that u_{jt} is a parameter that can be pre-computed using $\{\Omega_i\}_{i\in\mathcal{N}}$. We also note that x_{jt} belongs to $\mathcal{Z}_1$ if and only if $x_{jt} \le u_{jt}$, which means that (11) provides an alternative exact representation of $\mathcal{Z}_1$. This completes the proof.

□

2) Satisfying Demand: Satisfying demand is a critical operational requirement. We assume that the expected power demand $d_t \ge 0,\ t \in \mathcal{T}$ and the expected generation $g_t > 0,\ t\in\mathcal{T}$ are known quantities. To meet demand at time t, we need at least n_o operational WTs, where n_o must satisfy $n_o{\cdot}g_t \ge d_t$. As n_o is a non-negative integer, the latter condition is equivalent to $n_o \ge \left\lceil \frac{d_t}{g_t}\right\rceil$. Therefore, the maximum number of inactive WTs λ_t at time t must satisfy $J-\lambda_t = n_o \ge \left\lceil\frac{d_t}{g_t}\right\rceil$. In other words, $\lambda_t = J - \left\lceil\frac{d_t}{g_t}\right\rceil,\ t\in\mathcal{T}$.

A WT j can be inactive for one of two reasons either it has failed and has not yet been repaired or it is currently under maintenance. Thus, we can say that a WT is inactive if $\left(1-\sum_{k=1}^t x_{jt}\right)\cdot\mathbb{I}(\tilde{\omega}_j \le t) + x_{jt} = 1$. Using this formalism, meeting demand condition can be enforced by the following DRCC set:

$$\mathcal{Z}_2 = \Bigg\{\boldsymbol{X}\in\mathcal{D} : \inf_{\mathbb{P}\in\mathcal{P}^W(\delta)} \mathbb{P}\Bigg\{\tilde{\Omega} : \sum_{j\in\mathcal{J}}\left(1-\sum_{k=1}^t x_{jt}\right)\cdot\mathbb{I}(\tilde{\omega}_j\le t) + x_{jt} \le \lambda_t\Bigg\} \ge 1-\alpha, \quad t\in\mathcal{T}\Bigg\}.$$

Proposition 5. *Set $\mathcal{Z}_2$ admits an IP representation with following structure:*

$$\mathcal{Z}_2 = \left\{\boldsymbol{X}\in\mathcal{D} : \begin{array}{c} \frac{1}{N}\sum_{i\in\mathcal{N}} y_{it} \ge 1-\alpha, \quad t\in\mathcal{T} \\ \sum_{j\in\mathcal{J}}\left(1-\sum_{k=1}^t x_{jk}\right)\cdot\phi_{ijt} + x_{jt} \le \\ \lambda_t + M_{it}^2(1-y_{it}), \quad i\in\mathcal{N},\ t\in\mathcal{T} \\ y_{it}\in\{0,1\},\ i\in\mathcal{N},\ t\in\mathcal{T}\end{array}\right\}, \tag{13}$$

where $\{\phi_{ijt}\}_{i\in\mathcal{N},\,j\in\mathcal{J},\,t\in\mathcal{T}}$ are binary constants that can be pre-computed using $\{\Omega_i\}_{i\in\mathcal{N}}$.

Proof. Using Proposition 2, we obtain the following equality:

$$\mathcal{Z}_2 = \Bigg\{\boldsymbol{X}\in\mathcal{D} : \frac{1}{N}\sum_{i\in\mathcal{N}} \inf_{\Omega':\|\Omega'-\Omega_i'\|\le\delta} \mathbb{I}\left[\sum_{j\in\mathcal{J}}\left(\left(1-\sum_{k=1}^t x_{jk}\right)\cdot\mathbb{I}(\omega_j\le t) + x_{jt}\right)\le\lambda_t\right] \ge 1-\alpha,\ t\in\mathcal{T}\Bigg\}.$$

Using the fact that the indicator function is monotonous, we derive the following expression:

$$\inf_{\Omega':\|\Omega'-\Omega_i'\|\le\delta} \mathbb{I}\left\{\sum_{j\in\mathcal{J}}\left(\left(1-\sum_{k=1}^t x_{jk}\right)\cdot\mathbb{I}(\omega_j\le t)+x_{jt}\right)\le\lambda_t\right\} = \mathbb{I}\left\{\max_{\Omega':\|\Omega'-\Omega_i'\|\le\delta}\left\{\sum_{j\in\mathcal{J}}\left(\left(1-\sum_{k=1}^t x_{jk}\right)\cdot\mathbb{I}(\omega_j\le t)+x_{jt}\right)\right\}\le\lambda_t\right\}. \tag{14}$$

Since we are using the maximum norm, the indicator function of the right-hand side of (14) can be simplified as follows:

$$= \mathbb{I}\left\{\sum_{j\in\mathcal{J}}\left(\left(1-\sum_{k=1}^t x_{jk}\right)\cdot\max_{\omega_j:|\omega_j-\omega_j^i|\le\delta\cdot\sigma_j}\{\mathbb{I}(\omega_j\le t)\} + x_{jt}\right)\le\lambda_t\right\} \tag{15a}$$
$$= \mathbb{I}\left\{\sum_{j\in\mathcal{J}}\left(\left(1-\sum_{k=1}^t x_{jk}\right)\mathbb{I}(\max\{\omega_j^i-\delta\cdot\sigma_j,0\}\le t)+x_{jt}\right)\le\lambda_t\right\} \tag{15b}$$
$$= \mathbb{I}\left\{\sum_{j\in\mathcal{J}}\left(\left(1-\sum_{k=1}^t x_{jk}\right)\cdot\phi_{ijt}+x_{jt}\right)\le\lambda_t\right\}, \tag{15c}$$

where (15a) follows from the fact that the inner $\max()$ is separable in terms of j due to the maximum norm, (15b) holds due to the monotony of indicator function, and (15c) holds by the definition of ϕ_{ijt}, where $\phi_{ijt} = \mathbb{I}(\max\{\omega_j^i-\delta\cdot\sigma_j,0\}\le t),\ i\in\mathcal{N},\ j\in\mathcal{J},\ t\in\mathcal{T}$. Combining the previous results, we conclude that $\mathcal{Z}_2$ admits an IP representation with the following structure:

$$\mathcal{Z}_2 = \left\{\boldsymbol{X}\in\mathcal{D} : \begin{array}{c} \frac{1}{N}\sum_{i\in\mathcal{N}} y_{it} \ge 1-\alpha, \quad t\in\mathcal{T} \\ \sum_{j\in\mathcal{J}}\left(1-\sum_{k=1}^t x_{jk}\right)\phi_{ijt} + x_{jt} \le \\ \lambda_t + M_i^2(1-y_{it}), \quad i\in\mathcal{N},\ t\in\mathcal{T} \\ y_{it}\in\{0,1\},\ i\in\mathcal{N},\ t\in\mathcal{T}\end{array}\right\} \tag{16}$$

where $\{\phi_{ijt}\}_{i\in\mathcal{N},\,j\in\mathcal{J},\,t\in\mathcal{T}}$ are binary constants that can be pre-computed using $\{\Omega_i\}_{i\in\mathcal{N}}$.

□

Notice that in (16) we can easily find naive Big-M values by defining $M_{it}^2 = J-\lambda_t,\ i\in\mathcal{N},\ j\in\mathcal{J},\ t\in\mathcal{T}$. This definition is valid because $\sum_{j\in\mathcal{J}}\left(\left(1-\sum_{k=1}^t x_{jk}\right)\cdot\phi_{ijt}+x_{jt}\right)\le J$.

Remark 2. *Note that λ_t in* (16) *represents the maximum number of unavailable WTs while still satisfying power demand at time t. In practice, λ_t can be a random variable due to its dependence on power demand d_t and power generation g_t. Our model can accommodate these uncertainties, necessitating the minimization of λ_t with respect to d_t and p_t due to the monotonicity of the indicator operator in* (14)*. WTs may have different production profiles due to their distinct geographic position within the WF. To model this, it would be necessary to analyze $g_{jt}, j\in\mathcal{J}$ and redefine $\mathcal{Z}_2$ accounting for the different production profiles of each available WT. These two extensions will investigated in future works.*

3) Limiting the Number of Failures: We restrict the total number of corrective repairs within the planning horizon. To do this, we define the following DRCC set:

$$\mathcal{Z}_3 = \left\{ \boldsymbol{X} \in \mathcal{D} : \inf_{\mathbb{P} \in \mathcal{P}^W(\delta)} \mathbb{P}\left\{ \tilde{\Omega} : \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} \mathbb{I}(\tilde{\omega}_j \leq t) \cdot x_{jt} \leq \gamma \right\} \geq 1 - \beta \right\}. \tag{17}$$

Proposition 6. *Set $\mathcal{Z}_3$ admits an IP representation with following structure:*

$$\mathcal{Z}_3 = \left\{ \begin{array}{ll} \boldsymbol{X} \in \mathcal{D} : & \frac{1}{N} \sum_{i \in \mathcal{N}} w_i \geq 1 - \beta \\ & \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} \eta_{ijt} \cdot x_{jt} \leq \\ & \gamma + M_i^3 (1 - w_i), \quad i \in \mathcal{N} \\ & w_i \in \{0, 1\}, \ i \in \mathcal{N} \end{array} \right\}, \tag{18}$$

where $\{\eta_{ijt}\}_{i \in \mathcal{N},\, j \in \mathcal{J},\, t \in \mathcal{T}}$ are binary constants that can be pre-computed using $\{\Omega_i\}_{i \in \mathcal{N}}$.

Proof. The proof follows identical arguments to the proof of Proposition 5.

□

Notice that in (18), we can easily find a naive Big-M value by setting $M_i^3 = J - \gamma, \ i \in \mathcal{N}$. This definition is valid because $\sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} \eta_{ijt} \cdot x_{jt} \leq J, \ i \in \mathcal{N}$.

Corollary 1. *The proposed DRCC formulation in* (3) *can be reformulated exactly as a linear integer program.*

Proof. We note that the feasible region of the proposed DRCC program presented in (3) can be written as $\mathcal{X} = \mathcal{Z}_1 \cap \mathcal{Z}_2 \cap \mathcal{Z}_3$. Thus, using Prepositions 3, 4, 5, and 6, we conclude that $\mathcal{X}$ admits an exact IP representation.

□

E. Big-M Coefficients Strengthening

We have shown that $\mathcal{Z}_2$ and $\mathcal{Z}_3$ admit an IP representation, adopting the following common structure:

$$\mathcal{Z} = \left\{ x \in \{0,1\}^{n_x} : \begin{array}{l} \frac{1}{N} \sum_{i \in \mathcal{N}} z_i \geq 1 - p \\ a^{i\top} x \leq b^i + M^i (1 - z_i), i \in \mathcal{N} \\ z_i \in \{0, 1\}, \ i \in \mathcal{N} \end{array} \right\},$$

where $a^{i\top} \in \mathbb{Z}^{n_x}$ and $b^i \in \mathbb{Z}$ are non-negative sample-dependant parameters. Such sets represent a particular case of a finite scenario approximation of the chance-constrained binary packing problem investigated by [48]. Thus, we can utilize some of the results presented in [48] to strengthen the Big-M values used to represent $\mathcal{Z}_2$ and $\mathcal{Z}_3$. This leads to better relaxation, which speeds up the running times.

Specifically, we strengthen M^i using Algorithm 1. The maximization problem (20) of Line 2 corresponds to a continuous knapsack linear program, which can be solved by a simple sorting procedure – greedy algorithm. Algorithm 1 can be used to compute Big-M values for $\mathcal{Z}_2$ and $\mathcal{Z}_3$. For $\mathcal{Z}_3$, Algorithm 1 can be further improved by modifying Line 2 as follows:

$$\varphi_i(i') := \begin{array}{ll} \max & \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} \eta_{ijt} \cdot x_{jt} - \gamma \\ \text{s.t.} & \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} \eta_{i'jt} \cdot x_{jt} \leq \gamma \\ & \sum_{t \in \mathcal{T}} x_{jt} \leq 1, \ j \in \mathcal{J} \\ & x \in [0,1]^{n_x} \end{array} \tag{19}$$

We include the constraint $\sum_{t \in \mathcal{T}} x_{jt} \leq 1$, for each $j \in \mathcal{J}$. By doing this, we remove all infeasible solutions, such that for a given $j \in \mathcal{F}$, $x_{jt} = 1$ for more than one $t \in \mathcal{T}$, which are infeasible to our original optimization problem. As (19) includes more constraints than (20), it follows that the optimal value of (19) is smaller or equal to the optimal value of (20). Consequently, this modified algorithm may find lower Big-M values.

Next, we show that (19) can be solved by a simple ordering procedure in Proposition 7.

Proposition 7. *The optimization problem presented in* (19) *can be solved by a simple ordering procedure.*

Proof. See Appendix A.

□

Algorithm 1 Big-M Coefficients Strengthening [48]

1: **for** $i' \in \mathcal{N}$ **do**

$$\varphi_i(i') := \max\{a^{i\top} x - b^i : a^{i'\top} x \leq b^{i'}, x \in [0,1]^{n_x}\} \tag{20}$$

2: Sort $\{\varphi_i(i')\}_{i' \in \mathcal{N}}$ in a non-decreasing order:
$$\varphi_i(o_1) \leq \varphi_i(o_2) \leq \ldots \leq \varphi_i(o_N)$$

3: $M^i = \varphi_i(o_{\lfloor N \cdot p \rfloor + 1})$

IV. Computational Studies

We now discuss a computational study aimed at evaluating the performance of our optimization model. We present two case studies. In the first case study, we use real-world vibration monitoring data from a rotating machinery test rig to represent the degradation of a critical component in the WT. In the second case study, we use a simulation (bootstrapping) framework inspired by the same data obtained from the rotating machinery test rig to help evaluate the performance of our optimization model under various settings of data availability (and sparsity).

A. Computational Study with Real-world Degradation Signals

This case study was performed using a publicly available real-world vibration monitoring dataset obtained by performing accelerated degradation tests on rolling element bearings from an "as-good-as-new" state until failure. The actual test rig used to acquire the data has been described in detail in [49]. Note that bearings are a critical component of any WTs and are typically one of the key components that are continuously being monitored using vibration sensors. Hereafter, we will use the term *degradation signal* to refer to the vibration monitoring data, which evolves from the "as-good-as-new" state until the point of failure of the bearing. Bearings are assumed to have failed once their degradation signals cross a predefined failure threshold defined using ISO standards for vibration monitoring of industrial machinery [50].

1) Prognostics Modeling: We let $\mathcal{S}^m = \{S(0), ..., S(t_m)\}$ denote the m^{th} degradation signal, where $S(t)$ represents the amplitude of the degradation signal observed at time t We assume that a set of historical degradation signals (hereinafter referred to as *training data*) $\mathcal{S}^{\text{train}} = \{\mathcal{S}^1, \mathcal{S}^2, ...\}$ is available and can be used to estimate the parameters of the prognostic model. In this study, we consider 16 degradation signals that were split in half, 8 degradation signals used for training (estimating the parameters of the prognostic model), and 8 degradation signals used to test the optimization model. The 8 degradation signals used for testing were sampled with replacement in order to emulate the degradation of the WTs.

The prognostic model used in this study is similar to the one presented in [51]. The authors develop a stochastic model with a combination of deterministic and random parameters. They use a first-passage time approach to calculate the RLD. Specifically, they show that the RLD at time t can be approximated by an Inverse Gaussian distribution $IG(\nu_t, \gamma_t)$, where $\nu_t > 0$ is the location parameter and $\gamma_t > 0$ the shape parameter.

We use the *training data* to estimate the prior distributions of ν_t and γ_t. A Bayesian updating approach is then used to update the prior distributions using newly observed data as they become available. The updating process allows us to capture the latest degradation states of the WTs and update their RLDs accordingly. The updated RLDs are again used by the optimization model to compute a revised maintenance schedule.

2) Implementation of the Proposed DRCC Model: This continuous updating process of the RLDs motivates the implementation of the optimization model in a rolling horizon fashion with a freeze period of Δ^{upd} [days]. This means that only maintenance decisions scheduled within the freeze period $[1, ..., \Delta^{upd}]$ are performed. More specifically, it is assumed that when the maintenance crew is deployed in the WF, it only performs scheduled maintenance. This assumption is driven by the fact that performing maintenance on a set of WTs requires resources (e.g., personnel and spare parts) that need to be allocated beforehand. Therefore, the deployed maintenance crew is not allowed to change maintenance decisions. After the freeze period, the RLDs of WTs are re-estimated using fresh sensor data, and then the optimization model is re-executed leading to an updated maintenance schedule. This decision-updating process is repeated over and over so that the optimization model can adapt its decisions to the degradation process experienced by the components.

Our computational studies analyzed a WF with $J = 100$ WTs. The optimization model was implemented with following parameters: $T^{max} = 40$, $\Delta^{upd} = 20$, $c^{pr} = \$4K$, $c^{co} = 4c^{pr}$, $c^v = 12c^{pr}$, $p_t = \$0.4K$, $\rho_j = 7$, $\epsilon_j = 0.1$, $\lambda_t = 10$, $\alpha = 0.1$, $\gamma = 10$, and $\beta = 0.1$. The proposed DRCC formulation was tested for five different radius of the Wasserstein ball with $N = 100$. Additionally, two CBM benchmark models based on stochastic programming and robust optimization were implemented. It is worth recalling that the use of stochastic programming for optimizing WT maintenance schedule has been investigated in other papers [43], [44]. Thus, the stochastic programming model constitutes a valid state-of-the-art benchmark model. A summary of the implemented models and their corresponding notation used in the discussion of the numerical results is presented in Table I. Each simulation study included 20 updates of the rolling horizon, which is equivalent to 400 days of operation. The simulations were repeated 20 times for each configuration to account for the fluctuations in the simulated environment. The results obtained are displayed in Figure 2.

TABLE I: Summary of implemented models.

Category	Notation	Description
Proposed	DR: δ	Proposed DRCC strategy implemented with radius δ.
Benchmark	SAA	Stochastic programming formulation solved via Sample Average Approximation approach (SAA), which coincides with the proposed DRCC formulation with $\delta = 0$.
	Rob.	Robust optimization considering the worst-case scenario for each remaining lifetime, i.e., $\omega_j \in [\min_{i\in\mathcal{N}} \omega_j^i, \max_{i\in\mathcal{N}} \omega_j^i],\ j \in \mathcal{J}$.

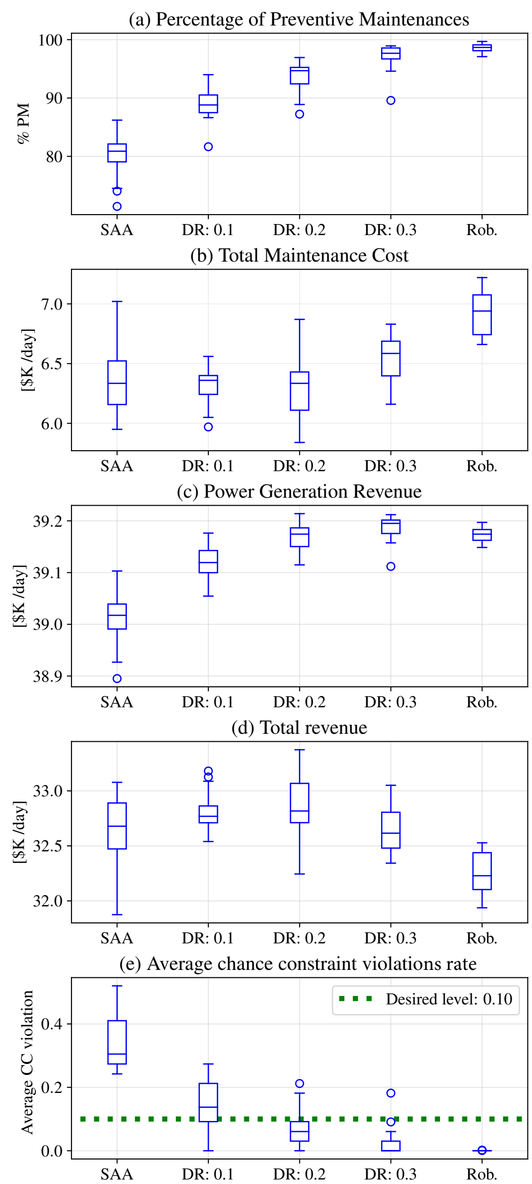

Fig. 2: Performance obtained with real-world degradation data.

3) Results: The results obtained are displayed in Figure 2. We notice that SAA model presents a deficient performance in terms of the percentage of preventive repairs, being close to 80%. Further, it can be seen that the chance constraint violation rate of SAA model exceeds the desired level–see Figure 2 (e). The poor performance of SAA model can be caused by the reduced number of training data sets or by the noise in the degradation signals, which negatively impacts the prediction accuracy of the prognostic model. This leads to a biased characterization of the model uncertainty, thereby making it invalid the implementation of stochastic programming in this setting. In contrast, Rob. model offers almost 100 % of preventive maintenances. However, it incurs high maintenance costs and thus is not competitive in terms of total revenue.

Our proposed DRCC model showcases significant improvements in terms of the percentage of preventive maintenance compared with SAA model. The highest total revenue of our model is reached by DR: 0.2. Nevertheless, the chance constraint violation probability of DR: 0.2 still exceeds the desired level so this alternative cannot be adopted. We select DR: 0.3, which yields a total revenue similar to SAA model but meets chance constraints requirements.

B. Computational Study with Synthetic Degradation Signals

This second case study relies on a simulation framework for generating CM data for rotating machinery. The simulation framework is inspired by a real-world vibration monitoring dataset used in [49] and is detailed in Appendix A and B of [39]. This case study also adopts the prognostic model presented by [51]. The optimization model was implemented using the same setting described in Section IV-A2.

We conducted two simulation studies to analyze the performance of the optimization model depending on the size of the training data (i.e., $|\mathcal{S}^{\text{train}}|$). These two settings are referred to as *sparse training data* ($|\mathcal{S}^{\text{train}}| = 5$) and *abundant training data* ($|\mathcal{S}^{\text{train}}| = 50$). The results obtained for these two settings are summarized in Figure 3.

1) Results: For the case of sparse training data (in red), it can be seen in Figure 3 that SAA model produces a poor performance, showing, on average, a percentage of preventive maintenance lower than 90 % and limited total revenue relative to the other models. More importantly, it is observed in Figure 3 (e) that the chance constraint violation rate of SAA model exceeds the desired level. This implies that operational requirements are not satisfied when using SAA model. This deficient performance is attributed to the fact that the prognostic model was trained with sparse data. As a consequence, the estimated RLDs may not accurately characterize the actual RLDs, thereby leading to biased maintenance decisions. Rob. model, on the other hand, presents an outstanding performance in terms of the percentage of preventive repairs with 100% for most cases. However, Rob. model is very conservative to accomplish this goal, which is reflected in the high total maintenance costs.

The performance of our proposed DRCC model varies depending on the value of δ. As discussed in Section III-A, δ captures the confidence level we have in the prognostic results. When training data is sparse, this confidence should be low. So we anticipate that increasing δ helps in seeking more robust maintenance decisions. This claim is supported by the computational results shown in Figure 3. We can see that increasing δ helps to raise the percentage of preventive maintenance and total revenue. For instance, with DR: 0.3, we obtain on average an increment of 6% in the percentage of preventive repairs and 0.9 % in the total revenue compared to SAA model. Furthermore, we can notice that DR: 0.3 satisfies the desired level of chance constraints. Therefore, our proposed DRCC model provides a reasonable balance between conservative and profitable decisions.

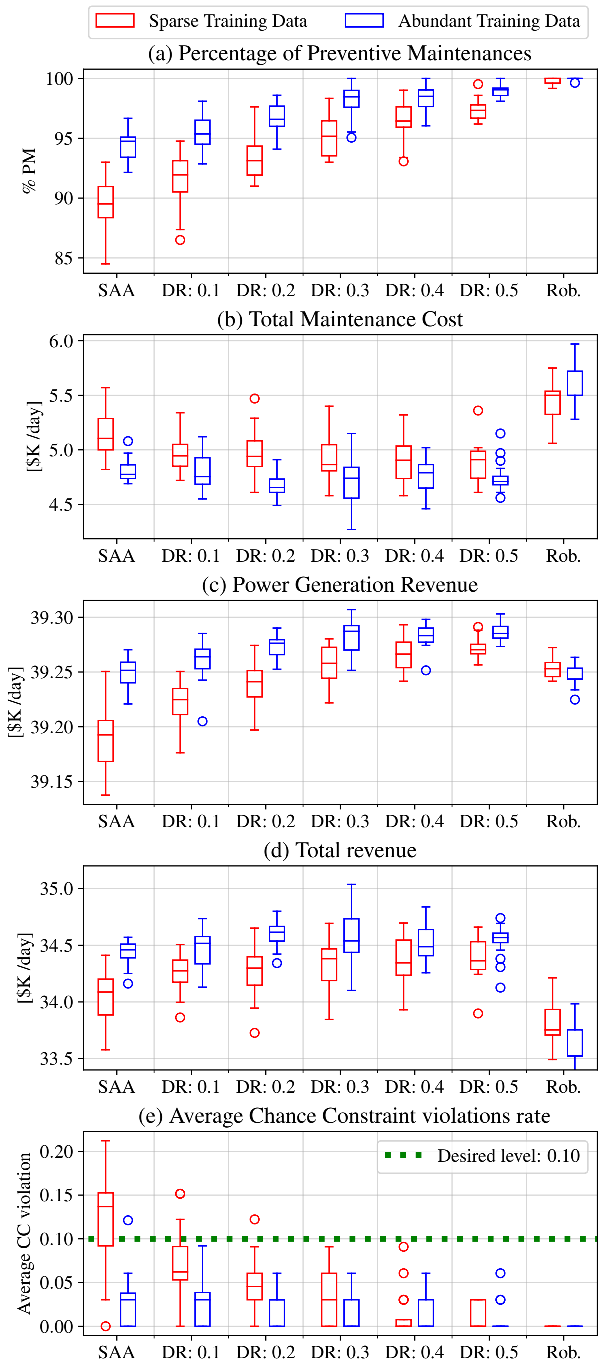

Fig. 3: Performance obtained with synthetic degradation data.

When training data is abundant (in blue), we can notice in Figure 3 that SAA model offers very competitive results. This is not surprising because in this setting it is expected to obtain accurate prognostic results, implying that the use of stochastic programming is well justified. Our proposed DRCC formulation also shows satisfactory results yet its improvement is lower compared with the case of sparse training data.

Remark 3. *One important step in implementing DRO models with Wasserstein ambiguity sets is determining a proper Wasserstein radius δ. In practical applications, the tuning of δ is accomplished through cross-validation [45], [46], relying on historical degradation sensor data. Specifically, cross-validation is used to evaluate the performance of the DRO maintenance strategy with different values of δ and then select the value that provides the best performance.*

C. Running Times

We evaluated the running times of the proposed DRCC formulation at different problem instances varying the number of WTs J. The optimization problems were solved on Python 3.10 using Gurobi 9.5. All the experiments were carried out on a state-of-the-art supercomputer–Partnership for an Advanced Computing Environment (PACE) with 48 GB of memory. Our formulation was implemented with two different methods to set the Big-M values of the integer representations of $\mathcal{Z}_2$ and $\mathcal{Z}_3$, denoted as: i) *Optimized Big-M* that uses the Big-M strengthening method of Section III-E and *Naive Big-M* that uses the naive Big-M values derived just after Propositions 5 and 6. Table II presents the maximum running times obtained in 50 executions for each configuration of J, respectively. We observe that Optimized Big-M implementation shows a significant reduction in running times compared with the naive values, being able to solve instances with 150 WTs in less than one hour.

TABLE II: Running times depending on the number of WTs ($N = 100$)

J	Optimized Big-M	Naive Big-M
	Max Running Time [s]	Max Running Time [s]
50	36.56	37.73
100	2375.25	4881.94
150	3590.80	> 10800

V. Conclusions

This paper developed the first CBM strategy for WTs that can theoretically deal with inaccurate prognostic results. The proposed CBM strategy relied on a DRCC formulation. The strategy used a prognostic model to estimate WTs' RLDs from sensor data. Then, the estimated RLDs were utilized to construct a Wasserstein ambiguity set, thereby capturing potential RLD perturbations. As a result, the DRCC formulation can compute reliable maintenance schedules even in presence of inaccurate prognostic results.

The effectiveness of our DRCC formulation was demonstrated in the computational studies involving real-world and synthetic degradation data, where the DRCC model outperformed the two benchmark models based on stochastic programming and robust optimization. The advantage of using our DRCC model was even more notable when dealing with inaccurate RLDs. In fact, when training data was sparse, our model produced on average an increment of 6% in the percentage of preventive repairs and 0.9 % in the total revenue compared to SAA model. In addition to these increments, our model was able to meet chance constraint requirements, which was not achieved by SAA model.

We proved that the proposed DRCC optimization problem can be reformulated as an integer linear program. We derived methods to strengthen the Big-M values of this reformulation, thereby enabling the DRCC model to be efficiently solved by off-the-shelf optimization software. We assessed the maximum running times of the DRCC optimization model, showing that the problem can be solved in less than one hour even in large instances with 150 WTs.

In future work, we plan to extend this model to account for uncertainty in wind power generation, power demand, and electricity prices. The objective will be to analyze the impact of these key factors on the optimal CBM schedule.

Appendix A

Proof of Proposition 7:

For any $j \in \mathcal{J}$ and $i, i' \in \mathcal{N}$, we can compute and sort the set $\{\eta_{ijt}/\eta_{i'jt}\}_{t\in\mathcal{T}}$, resulting the permutation $o^j(\cdot)$ that satisfies:

$$\frac{\eta_{ijo^j_1}}{\eta_{i'jo^j_1}} \le \frac{\eta_{ijo^j_2}}{\eta_{i'jo^j_j}} \le \ldots \le \frac{\eta_{ijo^j_{|\mathcal{T}}}}{\eta_{i'jo^j_{|\mathcal{T}|}}}.$$

We know that for $\mathcal{Z}_2$ it holds that $\eta_{ijt} \in \{0, 1\}$. Thus, some ratios may not be well-defined. When this happens, we use the following convention: $0/0 = 0$ and $1/0 = +\infty$.

Claim 1: There exists an optimal solution x^* to (19) such that for any $j \in \mathcal{J}$, $x^*_{jt} = 0,\ t \in \mathcal{T} \setminus \{o^j_{|\mathcal{T}|}\}$.

Proof. Let $\tilde{x}$ be an optimal solution to (19). Then, for a given $j \in \mathcal{J}$, we define three sets: $I^j_+ = \{t \in \mathcal{T} | \tilde{x}_{jt} > 0\}$, $T^{ij}_+ = \{t \in \mathcal{T} | \eta_{ijt} = 1\}$, and $T^{i'j}_+ = \{t \in \mathcal{T} | \eta_{i'jt} = 1\}$. We additionally define $b_j = \sum_{t\in\mathcal{T}} \eta_{i'jt} \cdot \tilde{x}_{jt}$. As $\tilde{x}$ is optimal to (19) , $\tilde{x}_{j,:}$ must be the optimal solution to the following two equivalent linear programs (LPs):

$$\begin{array}{lll} \max & \sum_{t\in\mathcal{T}} \eta_{ijt} \cdot x_{jt} & = \sum_{t\in T^{ij}_+} x_{jt} \\ \text{s.t.} & \sum_{t\in\mathcal{T}} \eta_{i'jt} \cdot x_{jt} \le b_j & = \sum_{t\in T^{i'j}_+ \cap T^{ij}_+} x_{jt} \le b_j \\ & \sum_{t\in\mathcal{T}} x_{jt} \le 1 & = \sum_{t\in T^{ij}_+} x_{jt} \le 1 \\ & x_{j,:} \in [0,1]^{|\mathcal{T}|} & x_{j,:} \in [0,1]^{|\mathcal{T}|} \end{array} \tag{21}$$

The optimal objective function value of (21) is lower or equal than 1. Then, for any $t \in I^j_+$ such that $t \neq o^j_{|\mathcal{T}|}$, we know that the following inequality holds:

$$\frac{\eta_{ijt}}{\eta_{i'jt}} \le \frac{\eta_{ijo^j_{|\mathcal{T}}}}{\eta_{i'jo^j_{|\mathcal{T}|}}}.$$

Therefore,

1) If $\eta_{ijo^j_{|\mathcal{T}|}} = 0$, it implies that all the other ratios are 0 and (21) becomes trivial, allowing $x_{jo^j_{|\mathcal{T}|}} = \min\{1, b_j\}$ and $x_{jt} = 0\ t \in \mathcal{T} - \{o^j_{|\mathcal{T}|}\}$ to be an optimal solution (21).
2) If $\eta_{i'jo^j_{|\mathcal{T}|}} = 0$, (21) is again trivial and its maximum is attained at $x_{jo^j_{|\mathcal{T}|}} = 1$ and $x_{jt} = 0\ t \in \mathcal{T} - \{o^j_{|\mathcal{T}|}\}$.
3) Else, $\eta_{ijo^j_{|\mathcal{T}|}} = 1$ and $\eta_{i'jo^j_{|\mathcal{T}|}} = 1$, which implies that $\eta_{i'jt} = 1,\ t \in T^{ij}_{+}$, thus resulting in $T^{ij}_{+} \subseteq T^{i'j}_{+}$. Subsequently, the constraints of (21) can be rewritten as $\sum_{t \in T^{ij}_{+}} x_{jt} \leq \min\{1, b_j\}$. The optimal value of the resulting LP can be attained at $x_{jo^j_{|\mathcal{T}|}} = \min\{1, b_j\}$ and $x_{jt} = 0,\ t \in \mathcal{T} \setminus \{o^j_{|\mathcal{T}|}\}$.

As the previous analysis is valid for any $\in \mathcal{J}$, it means that we can always compute an optimal solution to (19) that satisfies: $j \in \mathcal{J},\ x^*_{jt} = 0,\ t \in \mathcal{T} \setminus \{o^j_{|\mathcal{T}|}\}$. This finishes the proof of Claim 1. □

Claim 1 guarantees the optimal objective value of (19) does not change if we fix $x_{jt} = 0$ for any $j \in \mathcal{J}$ and $t \in \mathcal{T} \setminus \{o^j_{|\mathcal{T}|}\}$. This implies that we can remove these variables from the optimization program and still attain the same optimal value for the objective function. When removing these decision variables, we end up with the following optimization problem:

$$\begin{array}{ll} \max & \sum_{j \in \mathcal{J}} \eta_{ijo^j_{|\mathcal{T}|}} x_{jo^j_{|\mathcal{T}|}} - \gamma \\ \text{s.t.} & \sum_{j \in \mathcal{J}} \eta_{i'jo^j_{|\mathcal{T}|}} x_{jo^j_{|\mathcal{T}|}} \leq \gamma \ . \\ & x \in [0,1]^{n_x} \end{array} \tag{22}$$

This corresponds to a continuous knapsack problem that can be solved by a simple ordering procedure–greedy algorithm.

References

[1] R. Wiser, M. Bolinger, B. Hoen, D. Millstein, J. Rand, G. Barbose, N. Darghouth, W. Gorman, S. Jeong, and B. Paulos, "Land-based wind market report: 2022 edition," Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), Tech. Rep., 2022.

[2] M. Shafiee and J. D. Sørensen, "Maintenance optimization and inspection planning of wind energy assets: Models, methods and strategies," *Reliability Engineering & System Safety*, vol. 192, p. 105993, 2019.

[3] M. I. H. Tusar and B. R. Sarker, "Maintenance cost minimization models for offshore wind farms: A systematic and critical review," *International Journal of Energy Research*, vol. 46, no. 4, pp. 3739–3765, 2022.

[4] Z. Ren, A. S. Verma, Y. Li, J. J. Teuwen, and Z. Jiang, "Offshore wind turbine operations and maintenance: A state-of-the-art review," *Renewable and Sustainable Energy Reviews*, vol. 144, p. 110886, 2021.

[5] P. Papadopoulos, D. W. Coit, and A. A. Ezzat, "Seizing opportunity: Maintenance optimization in offshore wind farms considering accessibility, production, and crew dispatch," *IEEE Transactions on Sustainable Energy*, vol. 13, no. 1, pp. 111–121, 2021.

[6] F. Besnard, M. Patriksson, A.-B. Strömberg, A. Wojciechowski, K. Fischer, and L. Bertling, "A stochastic model for opportunistic maintenance planning of offshore wind farms," in *2011 IEEE Trondheim PowerTech*. IEEE, 2011, pp. 1–8.

[7] C. A. Irawan, D. Ouelhadj, D. Jones, M. Stålhane, and I. B. Sperstad, "Optimisation of maintenance routing and scheduling for offshore wind farms," *European Journal of Operational Research*, vol. 256, no. 1, pp. 76–89, 2017.

[8] Y. Zhou, J. Miao, B. Yan, and Z. Zhang, "Bio-objective long-term maintenance scheduling for wind turbines in multiple wind farms," *Renewable Energy*, vol. 160, pp. 1136–1147, 2020.

[9] B. Zhang and Z. Zhang, "A two-stage model for asynchronously scheduling offshore wind farm maintenance tasks and power productions," *International Journal of Electrical Power & Energy Systems*, vol. 130, p. 107013, 2021.

[10] P. Tavner, J. Xiang, and F. Spinato, "Reliability analysis for wind turbines," *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, vol. 10, no. 1, pp. 1–18, 2007.

[11] F. Spinato, P. J. Tavner, G. J. Van Bussel, and E. Koutoulakos, "Reliability of wind turbine subassemblies," *IET Renewable Power Generation*, vol. 3, no. 4, pp. 387–401, 2009.

[12] F. Ding and Z. Tian, "Opportunistic maintenance for wind farms considering multi-level imperfect maintenance thresholds," *Renewable Energy*, vol. 45, pp. 175–182, 2012.

[13] S. Carlos, A. Sánchez, S. Martorell, and I. Martón, "Onshore wind farms maintenance optimization using a stochastic model," *Mathematical and Computer Modelling*, vol. 57, no. 7-8, pp. 1884–1890, 2013.

[14] B. R. Sarker and T. I. Faiz, "Minimizing maintenance cost for offshore wind turbines following multi-level opportunistic preventive strategy," *Renewable energy*, vol. 85, pp. 104–113, 2016.

[15] H. Abdollahzadeh, K. Atashgar, and M. Abbasi, "Multi-objective opportunistic maintenance optimization of a wind farm considering limited number of maintenance groups," *Renewable Energy*, vol. 88, pp. 247–261, 2016.

[16] A. Erguido, A. C. Márquez, E. Castellano, and J. G. Fernández, "A dynamic opportunistic maintenance model to maximize energy-based availability while reducing the life cycle cost of wind farms," *Renewable Energy*, vol. 114, pp. 843–856, 2017.

[17] C. Zhang, W. Gao, S. Guo, Y. Li, and T. Yang, "Opportunistic maintenance for wind turbines considering imperfect, reliability-based maintenance," *Renewable energy*, vol. 103, pp. 606–612, 2017.

[18] C. Zhang, W. Gao, T. Yang, and S. Guo, "Opportunistic maintenance strategy for wind turbines considering weather conditions and spare parts inventory management," *Renewable Energy*, vol. 133, pp. 703–711, 2019.

[19] M. Li, M. Wang, J. Kang, L. Sun, and P. Jin, "An opportunistic maintenance strategy for offshore wind turbine system considering optimal maintenance intervals of subsystems," *Ocean Engineering*, vol. 216, p. 108067, 2020.

[20] L. Yang, G. Li, Z. Zhang, X. Ma, and Y. Zhao, "Operations & maintenance optimization of wind turbines integrating wind and aging information," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 1, pp. 211–221, 2020.

[21] J. McMorland, M. Collu, D. McMillan, J. Carroll, and A. Coraddu, "Opportunistic maintenance for offshore wind: A review and proposal of future framework," *Renewable and Sustainable Energy Reviews*, vol. 184, p. 113571, 2023.

[22] X. Zhu, J. Wang, and D. W. Coit, "Joint optimization of spare part supply and opportunistic condition-based maintenance for onshore wind farms considering maintenance route," *IEEE Transactions on Engineering Management*, 2022.

[23] C. A. Irawan, M. Eskandarpour, D. Ouelhadj, and D. Jones, "Simulation-based optimisation for stochastic maintenance routing in an offshore wind farm," *European Journal of Operational Research*, vol. 289, no. 3, pp. 912–926, 2021.

[24] J. Kang, J. Sobral, and C. G. Soares, "Review of condition-based maintenance strategies for offshore wind energy," *Journal of Marine Science and Application*, vol. 18, no. 1, pp. 1–16, 2019.

[25] E. Byon, L. Ntaimo, and Y. Ding, "Optimal maintenance strategies for wind turbine systems under stochastic weather conditions," *IEEE transactions on reliability*, vol. 59, no. 2, pp. 393–404, 2010.

[26] E. Byon and Y. Ding, "Season-dependent condition-based maintenance for a wind turbine using a partially observed markov decision process," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1823–1834, 2010.

[27] F. Besnard and L. Bertling, "An approach for condition-based maintenance optimization applied to wind turbine blades," *IEEE Transactions on Sustainable Energy*, vol. 1, no. 2, pp. 77–83, 2010.

[28] W. Liu, B. Tang, J. Han, X. Lu, N. Hu, and Z. He, "The structure healthy condition monitoring and fault diagnosis methods in wind turbines: A review," *Renewable and Sustainable Energy Reviews*, vol. 44, pp. 466–472, 2015.

[29] G. d. N. P. Leite, A. M. Araújo, and P. A. C. Rosas, "Prognostic techniques applied to maintenance of wind turbines: a concise and specific review," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1917–1925, 2018.

[30] M. Rezamand, M. Kordestani, R. Carriveau, D. S.-K. Ting, M. E. Orchard, and M. Saif, "Critical wind turbine components prognostics: A comprehensive review," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 12, pp. 9306–9328, 2020.

[31] Y. Peng, W. Qiao, F. Cheng, and L. Qu, “Wind turbine drivetrain gearbox fault diagnosis using information fusion on vibration and current signals,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021.

[32] Z. Tian, T. Jin, B. Wu, and F. Ding, “Condition based maintenance optimization for wind power generation systems under continuous monitoring,” *Renewable Energy*, vol. 36, no. 5, pp. 1502–1509, 2011.

[33] X. Lei and P. A. Sandborn, “Maintenance scheduling based on remaining useful life predictions for wind farms managed using power purchase agreements,” *Renewable Energy*, vol. 116, pp. 188–198, 2018.

[34] P. Zhou and P. Yin, “An opportunistic condition-based maintenance strategy for offshore wind farm based on predictive analytics,” *Renewable and Sustainable Energy Reviews*, vol. 109, pp. 1–9, 2019.

[35] T. Xia, Y. Dong, E. Pan, M. Zheng, H. Wang, and L. Xi, “Fleet-level opportunistic maintenance for large-scale wind farms integrating real-time prognostic updating,” *Renewable Energy*, vol. 163, pp. 1444–1454, 2021.

[36] Y. Lu, L. Sun, X. Zhang, F. Feng, J. Kang, and G. Fu, “Condition based maintenance optimization for offshore wind turbine considering opportunities based on neural network approach,” *Applied Ocean Research*, vol. 74, pp. 69–79, 2018.

[37] L. Yang, Y. Chen, and X. Ma, “A state-age-dependent opportunistic intelligent maintenance framework for wind turbines under dynamic wind conditions,” *IEEE Transactions on Industrial Informatics*, 2023.

[38] M. Yildirim, X. A. Sun, and N. Z. Gebraeel, “Sensor-driven condition-based generator maintenance scheduling—part i: Maintenance problem,” *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4253–4262, 2016.

[39] M. Yildirim, N. Z. Gebraeel, and X. A. Sun, “Integrated predictive analytics and optimization for opportunistic maintenance and operations in wind farms,” *IEEE Transactions on power systems*, vol. 32, no. 6, pp. 4319–4328, 2017.

[40] B. Basciftci, S. Ahmed, N. Z. Gebraeel, and M. Yildirim, “Stochastic optimization of maintenance and operations schedules under unexpected failures,” *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6755–6765, 2018.

[41] B. Basciftci, S. Ahmed, and N. Gebraeel, “Data-driven maintenance and operations scheduling in power systems under decision-dependent uncertainty,” *IISE transactions*, vol. 52, no. 6, pp. 589–602, 2020.

[42] I. Bakir, M. Yildirim, and E. Ursavas, “An integrated optimization framework for multi-component predictive analytics in wind farm operations & maintenance,” *Renewable and Sustainable Energy Reviews*, vol. 138, p. 110639, 2021.

[43] F. Fallahi, I. Bakir, M. Yildirim, and Z. Ye, “A chance-constrained optimization framework for wind farms to manage fleet-level availability in condition based maintenance and operations,” *Renewable and Sustainable Energy Reviews*, vol. 168, p. 112789, 2022.

[44] P. Papadopoulos, D. W. Coit, and A. Aziz Ezzat, “Stochos: Stochastic opportunistic maintenance scheduling for offshore wind farms,” *IISE Transactions*, pp. 1–15, 2022.

[45] W. Xie, “On distributionally robust chance constrained programs with wasserstein distance,” *Mathematical Programming*, vol. 186, no. 1, pp. 115–155, 2021.

[46] P. Mohajerin Esfahani and D. Kuhn, “Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations,” *Mathematical Programming*, vol. 171, no. 1-2, pp. 115–166, 2018.

[47] D. Bertsimas, S. Shtern, and B. Sturt, “A data-driven approach to multistage stochastic linear optimization,” *Management Science*, 2022.

[48] Y. Song, J. R. Luedtke, and S. Küçükyavuz, “Chance-constrained binary packing problems,” *INFORMS Journal on Computing*, vol. 26, no. 4, pp. 735–747, 2014.

[49] N. Z. Gebraeel, M. A. Lawley, R. Li, and J. K. Ryan, “Residual-life distributions from component degradation signals: A bayesian approach,” *IiE Transactions*, vol. 37, no. 6, pp. 543–557, 2005.

[50] W. S. Mitchell, M. P. Blake *et al.*, “Vibration and acoustic measurement handbook,” 1972.

[51] A. Elwany and N. Gebraeel, “Real-time estimation of mean remaining life using sensor-based degradation models,” *Journal of Manufacturing Science and Engineering*, vol. 131, no. 5, 2009. [Online]. Available: https://doi.org/10.1115/1.3159045

Heraldo Rozas received a B.S. and M.S. in electrical engineering from the University of Chile. He is currently pursuing a Ph.D. in Industrial Engineering at the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Tech. His research focuses on integrating data analytics and data-driven optimization to make more efficient decisions regarding the operation, maintenance schedule, and spare parts management in industrial applications. His work received the INFORMS ENRE Student Best Paper Award in 2023.

Weijun Xie is the Coca-Cola Foundation Early Career Professor and Assistant Professor in the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Tech. Dr. Xie obtained his Ph.D. in Operations Research at the Georgia Institute of Technology in 2017. His research interests are theory and applications of stochastic, discrete, and convex optimization. His works have received multiple awards, including the 2022 New Investigator Award from the Virginia Space Grant Consortium (NASA), the 2021 NSF CAREER Award, and the Winner of the 2020 INFORMS Young Researchers Paper Prize.

Nagi Gebraeel is the Georgia Power Early Career Professor and Professor in the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Tech. He received his MS and PhD from Purdue University in 1998 and 2003, respectively. His research interests lie at the intersection of IoT Predictive Analytics and optimization models for data-driven operational and logistical decision making. Dr. Gebraeel is a fellow of the Institute of Industrial and Systems Engineers (IISE), the former president of the IISE Quality and Reliability Engineering Division, and a member of the Institute for Operations Research and the Management Sciences (INFORMS).