# Data Fusion for Optimal Condition-Based Aircraft Fleet Maintenance With Predictive Analytics

**ZHENGYANG FAN**
**KUO-CHU CHANG**
**RAN JI**
**GENSHE CHEN**

**Maintaining and deploying a fleet of aircraft with limited resources and various mission requirements is both immensely challenging and of primary importance. Traditional preventive maintenance methods are static and inflexible and are not equipped to consider the complex dynamics of aircraft (e.g., wear and age), which may lead to low fleet availability and high maintenance costs. In this paper, we propose an integrated learn-then-optimize framework for condition-based predictive maintenance scheduling to support daily flight and maintenance planning by fusing data from multiple onboard sensors. The paradigm first predicts the remaining useful life for components of aircraft by using deep learning techniques, then models the fleet-level optimization as a constrained mixed-integer programming problem that captures different failure modes of aircraft and the available maintenance facilities. We also propose valid inequalities to improve the computational efficiency of the optimization model. Finally, we conduct a series of simulated experiments to validate the performance of the proposed predictive maintenance model. The numerical results show that the predictive maintenance model outperforms the traditional preventive maintenance model with respect to the mission accomplishment rate, aircraft availability rate, and cost effectiveness.**

Manuscript received October 26, 2022; revised November 5, 2023; released for publication April 4, 2024

Refereeing of this contribution was handled by Ramona Georgescu.

Z. Fan, K.-C. Chang, and R. Ji are with the Department of Systems Engineering and Operations Research, George Mason University, Fairfax, VA, 22030, USA (e-mail: zfan3@gmu.edu; kchang@gmu.edu; rji2@gmu.edu).

G. Chen is with the Intelligent Fusion Technology Inc., Germantown, MD, 20874, USA (e-mail: gchen@intfusiontech.com).

This work was supported by the Naval Air Warfare Center under Contract no. N68335-20-C-0306. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NAVAIR.

1557-6418/23/$17.00 

## I. INTRODUCTION AND RELATED WORK

Flight and maintenance planning (FMP) for military aircraft aims to identify optimal scheduling for a given fleet by (1) determining which aircraft are available to fly and for what duration, and (2) if and when to conduct maintenance on grounded aircraft. FMP aims to accomplish these mission tasks efficiently while also keeping operational costs at a minimum. FMP plays a critical role in guaranteeing the safety and reliability (e.g., fleet readiness rate and mission accomplishment rate) of military or commercial airlines.

Modern surface and aviation systems are designed with an ever-increasing level of automation and advanced machinery that include state-of-the-art sensors that monitor vital aircraft, ship, and auxiliary system functions. New tools and technologies are needed to augment current onboard condition monitoring and maintenance processes, improve system availability, increase operational readiness, and reduce life cycle costs. Along with the development of Industrial 4.0, which integrates sensors, software, and intelligent control to improve industrial processes, aircraft maintenance is transitioning from more traditional processes of corrective and preventive maintenance to a data-driven, predictive maintenance paradigm. While there have been significant strides made in utilizing machine learning (ML) and augmented intelligence (AI) for predictive maintenance, there is still a need to develop new tools that can produce more efficient and accurate condition-based predictive maintenance (CBPM) decisions [39]. Predictive maintenance involves analyzing machine data collected from various monitoring sensors, such as thermal, acoustic, vibration, pressure, and temperature, to generate meaningful insights about the machine's state, including failure classification, remaining useful life (RUL), and time to failure (TTF) [37]. The ultimate objective is to schedule proactive maintenance more accurately, enhance readiness, and improve efficiency in the logistics and supply chain. Such capability becomes especially crucial for mission-critical systems, ensuring sustained combat operations and readiness while minimizing costs and unplanned downtime.

The predictive maintenance approach fuses the data from on-board sensors to monitor the health condition of aircraft components to predict RUL prognostics of system and identify anomalous behavior, and thus turn equipment sensor data into meaningful, actionable insights for proactive maintenance in the anticipation of failure [39]. There are two main challenges in RUL-based predictive maintenance and flight scheduling problems. The first challenge is how to accurately predict the RUL prognostics for system components by exploiting the data from multiple sensors. The second challenge is the integration of RUL prognostics into FMP, considering the workforce capacity (e.g., availability of workstations and technicians to repair the components), flight mission requirements (e.g., type and

number of aircraft required to conduct different missions), and system reliability requirements.

For RUL prediction, most existing studies fall into three categories [33]: statistical-based models, conventional ML models, and deep learning models. Statistical-based models are built by fitting a probabilistic model to data by assuming that the degradation of system components over time can be characterized via an appropriate parametric function or a specified stochastic process model. For example, the Wiener process has been successfully employed to capture the degradation of components in bridge beams [31], thrust ball bearings [30], and micro-electron mechanical systems [35]. Gaussian process regression model is used to predict RUL prognostics of battery health [23] and slow-speed bearings [1]. The conventional ML algorithms, such as support vector machines, tree-based methods, and neural networks, have been extensively used in predictive maintenance in the past decades. For example, [28] developed a support vector regression model with a multi-class solver to identify various faulty patterns in rotating machines. Reference [17] constructed a random forest regression model to predict the RUL of spur gears. More recently, with the expansion of big data techniques, the popularity of deep learning algorithms for predictive maintenance has noticeably increased. Reference [12] developed recurrent neural networks (RNNs) for RUL prediction of bearings. Reference [34] designed a double convolutional neural network (CNN) architecture to predict RUL using time-series vibration signals. Reference [38] employed long short-term memory (LSTM) RNN to learn the long-term dependencies among degraded capacities and predict the RUL of lithium-ion batteries. We refer interested readers to [2] and [4] for a comprehensive survey of ML approaches, and to [27], particularly for deep learning approaches in predictive maintenance.

Numerous efforts have also been devoted to military aircraft fleet scheduling optimization. From the military perspective, one major concern in this problem is operational readiness [21], [24]. Reference [24] formulated a mixed-integer programming (MIP) model to maximize fleet availability under skilled workforce constraints. The model admits a network flow interpretation and can be solved efficiently by the branch-and-bound method. Reference [14] proposed a multiobjective MIP model to maximize fleet availability. To further improve the computational efficiency, [16] and [15] extended the work of [3] and developed heuristic algorithms to solve large-scale problem instances. Instead of directly maximizing fleet availability, [19] and [3] minimized the maximum number of aircraft in maintenance to be greater than the number of available maintenance spaces over the planning horizon. [21] introduced an MIP model for long-term planning of military aircraft by considering type-D heavy maintenance. Another major challenge for the FMP problem is computational efficiency. Some heuristic algorithms have been proposed to solve large-scale problem instances, for example, [8], [15], [16]. To obtain exact solutions with computational efficiency, [9] proposed an iterative algorithm that cuts off infeasible relaxation solutions via special valid inequalities. Reference [10] modify the classical $\epsilon$-constraint method to solve a biobjective quadratic program. More recently, [22] used ML models to predict the characteristics of optimal solutions, and added these characteristics to the original formulations to shrink solution space.

Though RUL prediction and maintenance scheduling optimization have received a considerable amount of attention from their own domains, very few studies have developed RUL prognostics and subsequently integrated the predicted RUL into maintenance scheduling. Reference [20] built an LSTM neural network to predict multiclass RUL prognostics for turbofan engines of aircraft, which are used subsequently to order and manage replacement spare parts. Reference [6] developed a particle filtering algorithm for RUL prediction of aircraft cooling units, and the predicted RUL was then passed to a linear programming optimization model to optimally schedule a fleet of aircraft maintenance considering spare parts. More recently, in the work of [18], the authors considered the maintenance of aircraft brakes using a threshold-based maintenance policy, i.e., once the predicted RUL falls below a user-defined threshold, the brake is replaced. They solved a multiobjective scheduling optimization model that seeks a trade-off between the minimization of flight delays, the number of unscheduled maintenance tasks, and the total number of maintenance tasks. Using predicted RUL as model coefficients, [32] proposed a multiobjective genetic algorithm for maintenance scheduling for a vehicle fleet by minimizing total cost, workload, and the expected number of failures and total changes in maintenance schedule. Most recently, [7] studied the alarm-based maintenance planning with imperfect RUL predictions for a fleet of vehicles by considering estimation uncertainties.

In this work, we propose an integrated learn-then-optimize framework for flight and maintenance scheduling with RUL predictions to maximize fleet-level operational availability and minimize costs. The proposed framework first employs advanced analytics from multiple onboard sensory data to draw meaningful insights to predict machine states and proactively schedule maintenance and flights to minimize costs and unplanned downtime. More specifically, we first develop a bidirectional long short-term memory (biLSTM) deep learning model to combine the time-series monitoring data for predicting the RUL of aircraft system components, and subsequently incorporate the RUL into an optimization model to determine the optimal fleet-level maintenance policies and flight scheduling by considering practical constraints such as workforce capacity and mission requirements. To the best of our knowledge, this paper represents the first study that explicitly considers three elements, i.e., (1) deep learning with multisensor data for RUL prediction, (2) predictive maintenance scheduling,

and (3) flight mission planning into an integrated learn-then-optimize framework. Our work differs from previous research in the sense that previous studies only focused on RUL prediction and predictive maintenance scheduling, while our proposed framework also takes into account the flight mission planning decision, workforce capacity constraints (e.g., different types and number of technicians required to conduct maintenance for different components of an aircraft system), and mission requirements (e.g., type and number of aircraft required to conduct specific missions), which are of particular importance to military applications.

The remainder of the paper is structured as follows. In Section II, we discuss the problem description and learn-then-optimize framework. Section III briefly introduces bidirectional LSTM and discusses the RUL prediction using multiple commercial modular aero-propulsion system simulation (C-MAPSS) engine sensory data. Section IV presents the MIP formulation for the FMP optimization model, followed by a set of valid inequalities to boost the computational speed. Various numerical experiments are conducted in Section V to demonstrate the superiority of the proposed predictive maintenance model over the traditional preventive maintenance method. In Section VI, we conclude our work.

## II. PROBLEM DESCRIPTION

Heavy equipment maintenance facilities such as aircraft service centers face the challenge of maximizing readiness while minimizing the costs of various maintenance tasks, subject to the availability constraints of specialty technicians, workstations, spare inventory, and mission requirements. FMP require making decisions about which aircraft should perform a mission and which aircraft should enter maintenance, while optimizing efficiency with limited resources, including technical workforce and maintenance workstations. Traditional maintenance scheduling is done via preventive maintenance, i.e., a fixed schedule to maintain each aircraft periodically to prevent failure. The proposed approach is to replace the preventive maintenance with a predictive maintenance strategy where the machinery conditions (such as RUL) are considered in order to proactively schedule the maintenance. The proposed approach is designed within a learn-then-optimize framework, in which we first apply deep learning to analyze the time-series monitoring sensory data for predicting equipment RUL and subsequently incorporate it into an optimization formulation to determine optimal fleet-level maintenance policies. In this section, we briefly introduce the elements of the proposed learn-then-optimize framework, covering the description of FMP problem setting, multicomponent aircraft system, and workforce capacity consideration.

### A. Problem Setting for FMP

We consider the problem of scheduling a set of predictive maintenance tasks with a given number of available maintenance stations and workforce capacity constraints. For each type of maintenance task, the skills required and the number of technicians with the skills needed to work on the task are assumed to be known (similar setting used in [24]). The tasks are to be performed at the available workstations by the needed number of technicians with a specific skill or multiple skills. All of the skilled technicians required for the set of tasks to be performed must be available in the given time frame.

The operational goal is to maximize the readiness of a fleet of aircraft with a certain mission requirement given the constraints. The key is to ensure sufficient availability of aircraft to meet the operational demands, which refer to all the flight activities (namely, waves or sorties) that are planned in a given period (say, 24 h). A mission flight of a single aircraft is called a sortie. More than one aircraft flying together is called a wave or mission.

Predictive planning is assumed to be executed daily based on the estimated RUL of each component for the aircraft. The 24-h planning horizon is assumed to begin at 6:00 pm (see [24]). The objective is to determine the schedule (i.e., the sequence in which the maintenance tasks are executed) with the maximum availability-to-cost ratio. Note that the optimal solution to this problem will yield a maintenance schedule in which (a) all mission tasks in the set will be performed as much as possible, (b) the available workforce and workstations will be utilized in an efficient manner, and (c) the minimum possible time to perform maintenance will be determined. Hence, a solution to this problem is critical to reducing maintenance downtime and maximizing fleet operational readiness.

### B. Multicomponent Aircraft System and Workforce Capacity

In this study, we consider the FMP problem with a fleet of aircraft, where each aircraft has a system of multiple repairable or replaceable components. Each aircraft is assumed to be of a specific type that can perform designated mission tasks. Each component of aircraft is assumed to fail independently of other components. Repair or replacement maintenance is scheduled based not only on the estimated RUL of this component in anticipation of a failure but also on the task's durational requirements. For example, suppose that the threshold value for triggering the replacement is when RUL reduces to 3 h. If a component has 4 h of RUL, then typically it will not be scheduled for maintenance and will be allowed to take on a task. However, if a mission task requires the flying duration to be 5 h, then this aircraft cannot fulfill this requirement and must either undergo

maintenance to replace the component or be assigned to fly other missions. In this study, we assume each aircraft has a multicomponent system, each component is independent from the others, and each component requires different technician specialties for repair or replacement. Generally, while flying is underway (and also immediately before and after), technicians are divided into three groups/specialties to perform the following activities:

- Trade 1: Weapons and armament electrical (WP)
- Trade 2: Airframe mechanical, airframe electrical and propulsion (AF)
- Trade 3: Avionics/electronics (AV)

During actual operations, each type of failure/maintenance requires several types/trades of technicians for service. For example, engine failure may require two AF technicians and three AV technicians to perform the repair, and radar failure may require one WP technician and one AV technician for repair. The explicit incorporation of mission requirements (e.g., type of aircraft and flying duration) as well as the workforce capacity (e.g., number and types of specialty technicians) increases the practicability and complexity of the FMP problem significantly.

## III. DEEP LEARNING FOR RUL PREDICTION

The accurate prediction of RUL prognostics provides important input to the scheduling optimization model. In this section, we will showcase our study on how to use LSTM deep learning approaches to predict RUL using a C-MAPSS engine degradation dataset [25].

### A. LSTM RNN

LSTM is one of the most widely used RNN architecture in time sequence ML modeling. It uses gates to control information flow in the recurrent computations and is excellent at holding long-term memories. Its gating mechanisms are ideally suitable for modeling the machinery degradation process [37]. The original version of LSTM was proposed in [13]. The popular Vanilla LSTM was introduced in [11], where forget gate was added to the LSTM architecture to improve the model. Figure 1 shows the vanilla LSTM cell [29]. LSTM memory cells consist of different neural networks, which are called gates. Gates are used to track the interactions between the memory units, and to decide which data should be remembered or forgotten during the training process. The input gate and output gate determine if the state memory cells can be modified by the input signal. The forget gate controls whether or not to forget the previous status of the signal.

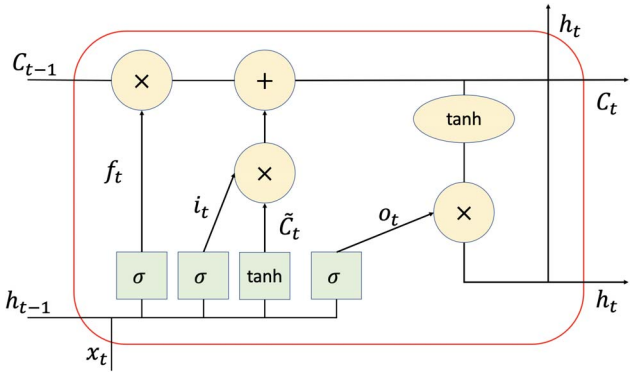



Figure 1. A vanilla LSTM cell [29].

In the cell, the functional relationships for each component are given as follows:

$$\begin{aligned} f_t &= \sigma\left(W_f \cdot x_t + R_f \cdot h_{t-1} + b_f\right) \\ i_t &= \sigma\left(W_i \cdot x_t + R_i \cdot h_{t-1} + b_i\right) \\ o_t &= \sigma\left(W_o \cdot x_t + R_o \cdot h_{t-1} + b_o\right) \\ \tilde{C}_t &= \varphi\left(W_c \cdot x_t + R_c \cdot h_{t-1} + b_c\right) \\ C_t &= f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \\ h_t &= o_t \cdot \phi\left(C_t\right), \end{aligned} \tag{1}$$

where $f_t, i_t$, and $o_t$ stand for forget gate, input gate, and output gate, respectively. Forget gate removes historical information from $C_{t-1}$; input and output gates control to update and output which part of information. $\sigma$, $\varphi$, and $\phi$ are nonlinear activation functions.

In order to obtain smooth states estimation of LSTM networks, bidirectional LSTM was proposed in [26], where a backward path is added to smooth out the prediction. Bidirectional LSTM thus can utilize the information in both forward and backward directions, which makes it suitable for intermediate prediction. Figure 2 describes the architecture of bidirectional LSTM.

Formulas for each component in backward path are almost identical to the forward LSTM model, which are

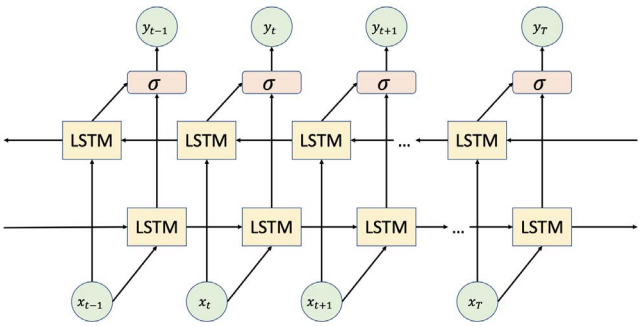



Figure 2. Bidirectional LSTM architecture.

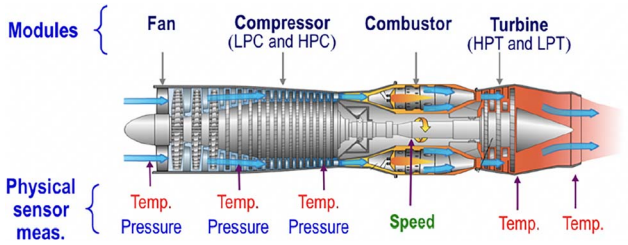



Figure 3. Turbofan engine model [37].

given below:

$$
\begin{aligned}
f'_t &= \sigma\left(W'_f \cdot x'_t + R'_f \cdot h'_{t-1} + b'_f\right) \\
i'_t &= \sigma\left(W'_i \cdot x'_t + R'_i \cdot h'_{t-1} + b'_i\right) \\
o'_t &= \sigma\left(W'_o \cdot x'_t + R'_o \cdot h_{t-1} + b'_o\right) \\
\tilde{C}'_t &= \varphi\left(W'_c \cdot x'_t + R'_c \cdot h'_{t-1} + b'_c\right) \\
C'_t &= f'_t \cdot C'_{t-1} + i'_t \cdot \tilde{C}'_t \\
h'_t &= o'_t \cdot \phi\left(C'_t\right),
\end{aligned}
\tag{2}
$$

where $f'_t, i'_t$, and $o'_t$ denote forget gate, input gate, and output gate, respectively (analogous to $f_t, i_t$ and $o_t$ as in LSTM model).

Table I
C-MAPSS Monitoring Sensor Data

| *Symbol* | *Description* | *Units* |
|---|---|---|
| T2 | Total temperature at fan inlet | R |
| T24 | Total temperature at LPC outlet | R |
| T30 | Total temperature at HPC outlet | R |
| T50 | Total temperature at LPT outlet | R |
| P2 | Pressure at fan inlet | psia |
| P15 | Total pressure in bypass-duct | psia |
| P30 | Total pressure at HPC outlet | psia |
| Nf | Physical fan speed | rpm |
| Ne | Physical core speed | rpm |
| epr | Engine pressure ratio (P50/P2) | – |
| Ps30 | Static pressure at HPC outlet | psia |
| phi | Ratio of fuel flow to Ps30 | pps/psi |
| NRf | Corrected fan speed | rpm |
| NRe | Corrected core speed | rpm |
| BPR | Bypass ratio | – |
| farB | Burner fuel-air ratio | – |
| htBleed | Bleed Enthalpy | – |
| Nf-dmd | Demanded fan speed | rpm |
| PCNfR-dmd | Demanded corrected fan speed | rpm |
| W31 | HPT coolant bleed | lbm/s |
| W32 | LPT coolant bleed | lbm/s |

## B. RUL Prognostics Prediction for Aircraft Engines

C-MAPSS is a tool to simulate performance of the turbofan engine, which is built under MATLAB and Simulink environment (see [25] for details). As shown

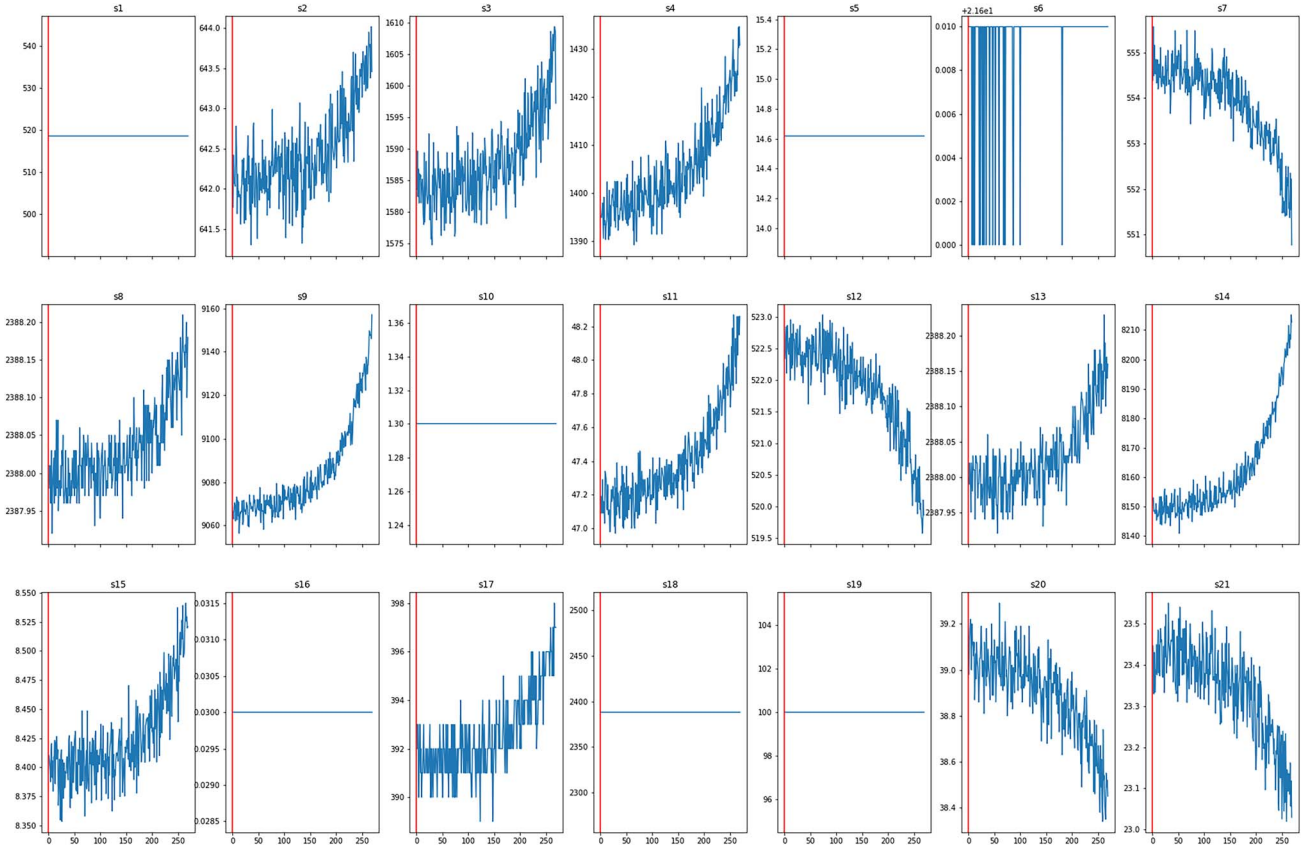



Figure 4. Run-to-failure sensor data examples.

in Fig. 3 (adapted from [37]), a turbofan engine typically consists of five modules: fan, low-pressure turbine (LPT), high-pressure turbine (HPT), low-pressure compressor (LPC) and high-pressure compressor (HPC).

The C-MAPSS dataset [25] includes hundreds of engine profiles with 21 onboard sensors monitoring the engine's health status (see Table I for descriptions of the sensors).

Three operation condition indicators (altitude, Mach numbers, and throttle resolver angle) are included as part of the observations as well. As an example, the C-MAPSS FD001 data set includes 200 engine profiles, 100 of which constitute a training set, where the historical run-to-failure measurement records are included. Figure 4 shows a run-to-failure data trajectories example from the 21 sensors under a specific operating condition. In the remaining testing set, sensor measurements are only recorded up to an early stage, and the goal is to predict the remaining engine life.

### C. LSTM Training and Testing

In order to train the LSTM model and predict RUL given the temporal degradation process in time-series data, we reshape input sensors' readings into blocks of size (30, 24), that is, a series of 30 consecutive sensor measurements (21 sensors) together with the operating condition indicator (3 indicators). For each input data block, the output is the corresponding remaining life at the last cycle in the time-series. To train a bidirectional LSTM model, the objective function is defined as the empirical mean squared error. Here, we select a two layer bidirectional LSTM with 145 neurons for each forward and backward layer in the network architecture. In the learning process, we apply the stochastic gradient descent (SGD) method with batch size 50 and learning rate of 0.0015. We test the RUL prediction performance with the C-MAPSS data. For example, Fig. 5 shows the comparison of true RUL and predicted RUL for testing dataset FD001. The root mean squared error (RMSE) obtained after testing 100 engines was found to be 15.16 operational cycles. This signifies that the average prediction error for the RUL of the engines is 15.16 operational cycles. Given that these results align with the best state-of-the-art algorithms' performance [36], this RMSE value is deemed a positive indicator of the predictive accuracy of our model. By fusing the 21 sensory time-series data with the bidirectional LSTM, the resulting RUL predictions are good indications of engine health condition and will be used to determine the optimal maintenance schedule.

## IV. PREDICTIVE MAINTENANCE AND MISSION SCHEDULING OPTIMIZATION

In this section, we first present our MIP formulation for FMP optimization model in Section IV-A, then in Section IV-B, we present two sets of valid inequalities to further improve the computational efficiency.

### A. Model Formulation

We introduce the notations used in the optimization model in Table II.

With these notations, the objective of the optimization model as defined in (3) is to maximize the normalized availability. We define normalized availability as the ratio of the weighted sum of available aircraft over the total maintenance cost over the planning horizon,

$$\max \quad \frac{\sum_{t=1}^{T}\sum_{k=1}^{K} M_t^k \cdot \left(\sum_{i=1}^{A_k}\left(1 - z_{it}^k\right)\right)}{\sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{i=1}^{A_k}\sum_{f=1}^{F} c_f^k x_{itf}^k}. \tag{3}$$

The weight values indicate the relative importance of mission in a particular operational cycle; thus, higher weights are assigned to more critical missions. The denominator represents the total maintenance cost in the planning horizon. The optimal maintenance schedules

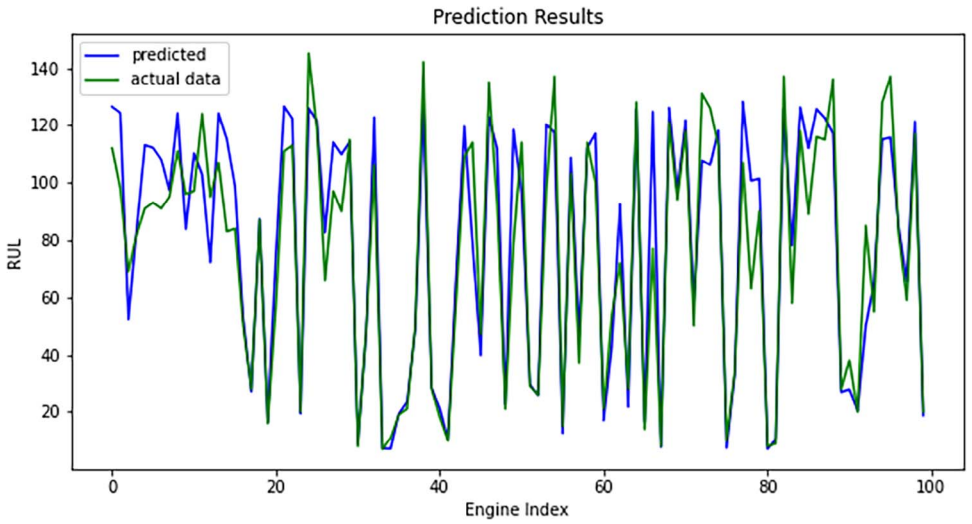



Figure 5. RUL prediction results on FD001 dataset.

Table II
Notation

| | |
|---|---|
| **Sets**: | |
| $T$ | Planning horizon ($t = 1, 2, ..., T$) |
| $K$ | Type of aircraft ($k = 1, 2, ..., K$) |
| $F$ | Type of failure mode/component ($f = 1, 2, ..., F$) |
| $R$ | Type of technicians ($r = 1, 2, ..., R$) |
| **Parameters**: | |
| $S$ | Number of the maintenance station |
| $A_k$ | Number of type-$k$ aircraft |
| $M_t^k$ | Number of required aircraft of type $k$ conducting missions at time $t$ |
| $D_t^k$ | Mission duration of aircraft type $k$ at time $t$ |
| $I_{fr}^k$ | Number of technicians of trade $r$ required to rectify the failure mode $f$ for aircraft type $k$ |
| $E_{fr}^k$ | Maintenance time of trade $r$ required to rectify the failure mode $f$ for aircraft type $k$ |
| $\lambda_r$ | Number of technicians of trade $r$ available at the initial period of the planning horizon |
| $r_{if}^k$ | Initial RUL for component $f$ of aircraft $i$ of type $k$ |
| $l_{tf}^k$ | Duration of mission for component $f$ of type $k$ aircraft at time $t$ |
| $c_f^k$ | Fixed charge cost for repairing component/failure mode $f$ for one aircraft of type $k$ |
| $E_{i0fr}^k$ | Remaining maintenance time for trade $r$ to repair aircraft $i$ of type $k$ |
| $r_{max,f}^k$ | RUL for component $f$ of aircraft type $k$ after maintenance (i.e., the maximum RUL) |
| **Decision Variables**: | |
| $x_{itf}^k$ | $= \begin{cases} 1 & \text{if aircraft } i \text{ of type } k \text{ enters maintenance station in time } t \text{ to rectify failure mode } f \\ 0 & \text{otherwise} \end{cases}$ |
| $z_{itfr}^k$ | $= \begin{cases} 1 & \text{if aircraft } i \text{ of type } k \text{ is in maintenance in time } t \text{ to rectify failure mode } f \text{ by trade } r \\ 0 & \text{otherwise} \end{cases}$ |
| $z_{itf}^k$ | $= \begin{cases} 1 & \text{if aircraft } i \text{ of type } k \text{ is in maintenance in time } t \text{ to rectify failure mode } f \\ 0 & \text{otherwise} \end{cases}$ |
| $z_{it}^k$ | $= \begin{cases} 1 & \text{if aircraft } i \text{ of type } k \text{ is in maintenance in time } t \\ 0 & \text{otherwise} \end{cases}$ |
| $y_{it}^k$ | $= \begin{cases} 1 & \text{if aircraft } i \text{ of type } k \text{ conducts mission in time } t \\ 0 & \text{otherwise} \end{cases}$ |
| $r_{itf}^k$ | RUL of component/failure mode $f$ for aircraft $i$ of type $k$ in time $t$ |

are to be obtained subject to a number of constraints, as shown below.

Mission Requirement Constraint

$$\sum_{i=1}^{A_k} y_{it}^k = M_t^k \qquad t \in T, k \in K. \quad (4)$$

Constraint (4) enforces the demands of mission requirements to be satisfied for aircraft type $k$ in each period $t$. That is, the sum of all assigned missions needs to satisfy the mission requirement.

RUL Dynamic Constraints

$$r_{i1f}^k = r_{if}^k \qquad \forall i \in I, k \in K, f \in F, \quad (5)$$

$$r_{i,t+1,f}^k \le r_{itf}^k - y_{it}^k \cdot l_{tf}^k + r_{\max,\mathrm{f}}^k \cdot x_{itf}^k \qquad t = 1, ..., T-1, k \in K, i \in I, f \in F, \quad (6)$$

$$r_{i,t+1,f}^k \ge r_{itf}^k - y_{it}^k \cdot l_{tf}^k \qquad t = 1, ..., T, k \in K, i \in I, f \in F, \quad (7)$$

$$r_{i,t+1,f}^k \le r_{\max,\mathrm{f}}^k \qquad t = 1, ..., T-1, k \in K, i \in I, f \in F, \quad (8)$$

$$r_{i,t+1,f}^k \ge r_{\max,\mathrm{f}}^k \cdot x_{itf}^k \qquad t = 1, ..., T-1, k \in K, i \in I, f \in F, \quad (9)$$

Constraints (5)–(9) model the behavior of RUL in failure mode $f$ for aircraft $i$ of type $k$ at time $t$. In other words, these constraints model the dynamics of components' RUL when a mission assignment or maintenance activity is conducted. Note that our assumption is that different components of an aircraft may exhibit different rates of degradation when a mission is conducted, resulting in varying reductions in the remaining operational cycles for each component, thus leading to different durations of missions for different components. For instance, when an aircraft undertakes a mission with a duration of 2 h, its weapon-related components may reduce their RUL by one operational cycle, while electrical and propulsion components may reduce their RUL by two operational cycles. However, in our numerical experiment section, we've simplified this by assuming a global reduction of RUL for all components, implying that all components experience the same RUL reduction.

Maintenance State Dynamic Constraints

$$x^k_{itf} \le z^k_{i,t+y,f,r} \qquad t = E^k_{i0fr}+1, ..., T - E^k_{fr}, \; y = 1, ..., E^k_{fr}, i \in I, k \in K, r \in R, \quad (10)$$

$$m \cdot x^k_{itf} \ge z^k_{i,t+y,f,r} \qquad t = E^k_{i0fr}+1, ..., T - E^k_{fr}, \; y = 1, ..., E^k_{fr}, i \in I, k \in K, r \in R, \quad (11)$$

$$x^k_{itf} + x^k_{i,t+y,f} \le 1 \qquad t = E^k_{i0fr}+1, ..., T - E^k_{fr}, \; y = 1, ..., E^k_{fr}, i \in I, k \in K, r \in R, \quad (12)$$

$$x^k_{itf} + x^k_{i,t+y,f} \le 1 \qquad t = T - E^k_{fr}+1, ..., T, \; y = 1, ..., T-t, i \in I, k \in K, r \in R, \quad (13)$$

$$m \cdot x^k_{itf} \ge z^k_{i,t+y,fr} \qquad t = T - E^k_{fr}+1, ..., T, \; y = 1, ..., T-t, i \in I, k \in K, r \in R, \quad (14)$$

$$x^k_{itf} + x^k_{i,t+y,f} \le 1 \qquad t = T - E^k_{fr}+1, ..., T, \; y = 1, ..., T-t, i \in I, k \in K, r \in R, \quad (15)$$

$$z^k_{itfr} = 1 \qquad t = 1, ..., E^k_{i0fr}, i \in I, \; k \in K, r \in R, f \in F, \quad (16)$$

$$x^k_{itf} + z^k_{itfr} \le 1 \qquad t = 1, ..., E^k_{i0fr}, i \in I, \; k \in K, r \in R, f \in F, \quad (17)$$

$$z^k_{itf} \ge z^k_{itfr} \qquad t = 1, ..., E^k_{i0fr}, i \in I, k \in K, r \in R, f \in F, \quad (18)$$

$$z^k_{itf} \le \sum_{r=1}^{R} z^k_{itfr} \qquad t \in T, k \in K, i \in I, f \in F, \quad (19)$$

$$z^k_{it} \ge z^k_{itf} \qquad t \in T, k \in K, i \in I, f \in F, \quad (20)$$

$$z^k_{it} \le \sum_{f=1}^{F} z^k_{itf} \qquad t \in T, k \in K, i \in I, \quad (21)$$

$$z^k_{it} + y^k_{it} \le 1 \qquad t \in T, k \in K, i \in I. \quad (22)$$

The maintenance state dynamic constraints are a set of logic constraints: When an aircraft performs its maintenance activity, it will remain in the maintenance state and cannot participate in other activities, such as flight missions. Constraints (10)–(17) model relationship between $x^k_{itf}$ and $z^k_{itfr}$ for different time segments. Specifically, constraints (16) and (17) enforce that for aircraft that do not complete maintenance during the previous planning horizon (the previous day), they will stay in maintenance and will not be available until $\max_r E^k_{i0fr}$. Notice that constraints (10)–(17) consist of two parts: The first part includes constraints (16) and (17), which are meaningful when there are unfinished maintenance left from the previous day, e.g., $E^k_{i0fr} > 0$. These constraints model the dynamics for the periods until the leftover maintenance is completed. The second part, constraints (10)–(15), models the maintenance state dynamics after the completion of unfinished maintenance from the previous day. Constraints (18)–(21) impose relationships among $z^k_{itfr}$, $z^k_{itf}$, $z^k_{it}$. For example, if, for some $r$, $z^k_{itfr} = 1$, then, constraint (18) will ensure $z^k_{itf} = 1$, and constraint (20) will force $z^k_{it} = 1$. That means, if some trade $r$ is working on an aircraft, then this aircraft is in maintenance at time $t$. Constraint (22) states that if an aircraft is in maintenance, it is not available for any missions at that time.

Resource and Mission Conflict Constraints

$$\sum_{k=1}^{K}\sum_{i=1}^{A_k} z^k_{it} \le S \qquad t \in T, \quad (23)$$

$$\sum_{k=1}^{K}\sum_{i=1}^{A_k}\sum_{f=1}^{F} z^k_{itfr} \cdot I^k_{fr} \le \lambda_r \qquad t \in T, r \in R, \quad (24)$$

$$y^k_{i,t+y} + y^k_{it} \le 1 \qquad t \in T, k \in K, i \in I, y = 1, ..., D^k_t, \quad (25)$$

$$y^k_{it} + x^k_{i,t+y,f} \le 1 \qquad t \in T, k \in K, i \in I, y = 1, ..., D^k_t. \quad (26)$$

Constraints (23) and (24) are resource constraints: Constraint (23) requires that one can at most maintain $S$ (the number of maintenance stations we have) aircraft simultaneously. Constraint (24) describes that the required number of maintenance technicians cannot exceed the number of available technicians. Constraint (25) enforces that an aircraft can only conduct one mission at a time. Constraint (26) prevents an aircraft that is in maintenance from conducting a mission.

B. Reformulation and Valid Inequality

1) Reformulation of Objective Function: Notice that the objective function (3) is highly non-linear. We convert it to a linear objective function via epigraphic reformulation, namely,

$$\max \quad \lambda \quad (27)$$

with a new constraint involving bilinear terms

$$\lambda \sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{i=1}^{A_k}\sum_{f=1}^{F} c^k_f \cdot x^k_{itf} \le \sum_{t=1}^{T}\sum_{k=1}^{K} M^k_t \cdot \left(\sum_{i=1}^{A_k}\left(1 - z^k_{it}\right)\right).$$

Since bilinear programming problems are notoriously difficult to solve in practice, by introducing several

auxiliary variables and a number of constraints, we further linearize the above bilinear constraint via McCormick inequalities:

$$w_{ift}^k \le \lambda \qquad t = 1, ..., T, k \in K, i \in I, f \in F, \quad (28)$$

$$w_{ift}^k \le m \cdot x_{ift}^k \qquad t = 1, ..., T, k \in K, i \in I, f \in F, \quad (29)$$

$$w_{ift}^k \ge -m \cdot \left(1 - x_{ift}^k\right) + \lambda \qquad t = 1, ..., T, k \in K, i \in I, f \in F, \quad (30)$$

$$\sum_{t=1}^{T}\sum_{k=1}^{K} M_t^k \cdot \left(\sum_{i=1}^{A_k}\left(1 - z_{it}^k\right)\right) \ge \sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{i=1}^{A_k}\sum_{f=1}^{F} c_f^k w_{itf}^k. \quad (31)$$

Finally, with the transformed objective function (27) and the constraints (4)–(26) and (28)–(31), we obtain a mixed-integer linear programming formulation for our CBPM scheduling problem, which can be solved efficiently by using commercial optimization solvers like CPLEX or GUROBI. Note that in reality, some of the desired missions may not be fulfilled even when all aircraft are in good condition. We therefore convert the mission requirement constraint (4) into an inequality constraint by allowing part of the missions to not be completed.

2) Valid Inequalities: Although the MIP formulation proposed in Section III-A can be directly solved by using commercial solvers such as GUROBI and CPLEX, it can be time-consuming for some parameter instances. Based on some preliminary computational experiments, we noticed that solvers may not obtain optimal integer solutions within a reasonable amount of time when mission durations are small and mission requirements are dense (always have mission during the day shift). We therefore developed two classes of valid inequalities to improve computational efficiency. In order to derive valid inequalities, we further assume that each component of each aircraft only needs to be maintained at most once in the planning horizon and that it is available for mission assignment after the maintenance. This assumption is not restrictive since we are planning over a daily horizon, and we mainly focus on maintenance types such as line maintenance or line-replaceable units (LRUs) replacement. Before we formally describe the valid inequalities for solving the proposed MIP model, we first review $0-1$ knapsack set and cover inequality from integer programming.

Given $b > 0$ and $a_j > 0$ for $j \in N := \{1, 2, \ldots, n\}$, the $0 - 1$ knapsack set is defined as

$$K := \left\{ x \in \{0, 1\}^n : \sum_{j=1}^{n} a_j x_j \le b \right\}.$$

A set $C \subseteq N$ is called a cover if $\sum_{j \in C} a_j > b$, and a cover inequality corresponding to the cover $C$ is given by

$$\sum_{j \in C} x_j \le |C| - 1, \quad (32)$$

where $|C|$ denotes the cardinality of set $C$. The cover inequality (32) is a valid inequality for set $K$. Readers are refereed to [5] for more details about knapsack inequality.

**Proposition 1.** *For each component of each aircraft of each aircraft type, the inequality*

$$\sum_{t \in C} y_{it}^k \le |C| - 1 + \sum_{t=1}^{c_0 - 1} x_{itf}^k \quad (33)$$

*is valid for original MIP formulation described by constraints (4)–(26), where C is the cover of the following knapsack constraint:*

$$\sum_{t=1}^{24} D_t^k y_{it}^k \le r_{if}^k, \quad (34)$$

*and $c_0$ denotes the smallest element of cover C.*

**Proof.** We will consider two cases:

Case-1. $\sum_{t=1}^{c_0-1} x_{itf}^k = 0$: In this case, $i$th aircraft of type $k$ aircraft will not be maintained during time $t = 1$ to time $t = c_0 - 1$. The inequality is reduced to cover inequality

$$\sum_{t \in C} y_{it}^k \le |C| - 1,$$

which is valid since we cannot let the aircraft conduct all missions at time $t \in C$ without maintaining the aircraft to increase its RUL $r_{if}^k$.

Case-2. $\sum_{t=1}^{c_0-1} x_{itf}^k = 1$: In this case, the aircraft is maintained at some time point between time $t = 1$ and time $t = c_0 - 1$. Then the inequality reduces to

$$\sum_{t \in C} y_{it}^k \le |C|,$$

which is redundant since, by our assumption, the aircraft can conduct any mission after the maintenance. □

Notice that we cannot identify all possible cover inequalities in general, and a separation algorithm is needed to find violated cover inequalities. In our case, however, we can directly identify all cover inequalities by enumeration since we only have 24 variables in knapsack constraint

$$\sum_{t=1}^{24} D_t^k y_{it}^k \le r_{if}^k.$$

Also, the aircraft are usually scheduled during the day shift (starts at 6:00 am and ends at 6:00 pm) [24]; we thus only need to identify all possible cover $C$ of knapsack

constraint

$$\sum_{t=13}^{24} D_t^k y_{it}^k \le r_{if}^k.$$

This is because on the night shift, the mission duration parameter $D_t^k$ will be 0. Therefore, for each component of each aircraft (i.e., for each $i$, $k$, and $f$), we only need to enumerate $2^{12} = 4096$ cover inequalities. However, when the planning horizon is longer (e.g., more than 24 h), enumerating all cover inequalities can be extremely time-consuming since the number of required inequalities is exponential in the length of the planning horizon. In our case, according to our numerical experiment, it will take less than 30 s to find all covers when we have a fleet of 12 aircraft with a planning horizon of 1 day.

The valid inequalities presented in Proposition 1 are usually not strong, hence insufficient to solve large-scale practical problems. Our next result shows a way to strengthen the valid inequalities in Proposition 1.

**Proposition 2.** *For each component of each aircraft of each aircraft type, the inequality*

$$\sum_{t\in C} y_{it}^k + \sum_{t\in C_1} y_{it}^k \le |C| - 1 + (1 + |C_1|) \sum_{t=1}^{c_0-1} x_{itf}^k$$

*is valid for original MIP formulation described by constraints (4)–(26), where $C$ is the cover of the following knapsack constraint*

$$\sum_{t=1}^{24} D_t^k y_{it}^k \le r_{if}^k,$$

*and $c_0$ denotes the smallest element of cover $C$. $C_1$ is defined as*

$$C_1 := \{t' : c_1 + 1 \le t' \le 24, D_t^k \ge \max_{t\in C} D_t^k\},$$

*where $c_1$ denotes the largest element of cover $C$.*

**Proof.** Similar to the proof of Proposition 1, we will consider two cases:

Case-1. $\sum_{t=1}^{c_0-1} x_{itf}^k = 0$: In this case, $i$th aircraft of type $k$ aircraft will not be maintained during time $t = 1$ to time $t = c_0 - 1$. The inequality is reduced to cover inequality

$$\sum_{t\in C} y_{it}^k + \sum_{t\in C_1} y_{it}^k \le |C| - 1,$$

which is an extended cover inequality of cover inequality

$$\sum_{t\in C} y_{it}^k \le |C| - 1,$$

hence valid for original formulations.

Case-2. $\sum_{t=1}^{c_0-1} x_{itf}^k = 1$: In this case, the aircraft is maintained at some time point between time $t = 1$ and time $t = c_0 - 1$. Then the inequality reduces to

$$\sum_{t\in C} y_{it}^k + \sum_{t\in C_1} y_{it}^k \le |C| + |C_1|,$$

which is redundant, hence valid for original formulations. □

## V. NUMERICAL EXPERIMENT

To validate the proposed model and solution method, we conduct a series of numerical experiments to illustrate the model performance and computational efficiency. In Section V-A, we explain the experimental designs and parameter settings for the computational tests. Section V-B shows the performance of the proposed predictive maintenance model compared to the traditional preventive maintenance model. Section V-C evaluates the impact of two classes of valid inequalities (introduced in Section IV-B) on computational efficiency. The experiments are conducted on a computer with an Intel Core i7 4-core 2.7 GHz processor and 16 GB of RAM. All mathematical programs are coded in Python 3 language, and all problem instances are solved by GUROBI 9.1 under Python API.

### A. Experimental Design and Parameters Setting

Before we provide a formal presentation and interpretation of our numerical results, we will first introduce the workflow of the learn-then-optimize framework that we used to generate the results. Given a set of onboard sensor readings, we first predict RUL for different components of an aircraft by fusing the data with the bidirectional LSTM neural network model (introduced in Section III). The values of the initial estimated RUL in Table V are obtained based on these predictions. The predicted RUL is then used as input parameters for the optimization model (proposed in Section IV). Together with other parameters such as mission and maintenance requirements, we run our optimization model and finally obtain the scheduling decisions.

The parameter settings for our experiment are given as follows: Planning horizon $T = 24$ h (since we are planning daily operations), types of aircraft $K = 2$, with 6 aircraft for each type, i.e., a total of 12 aircraft. For each type of aircraft, there are three different failure modes to be considered, e.g., WP, AF, and AV. Each failure mode requires several technicians to repair it. The number of maintenance stations is set to be three, so no more than three aircraft can be maintained at the same time. The number of available technicians for each trade (i.e., trades 1, 2, and 3) at the beginning of the planning horizon is set to be 10.

For the experiment scenario, Table III summarizes the mission type (i.e., the number and type of aircraft $M_t^k$ and mission during time $D_t^k$) for a 24-h horizon. In this example, all mission types in the first 12 h (6:00 pm to 6:00 am) are the same, during which there are no

Table III
Mission Requirements for Number and Type of Aircraft $M_t^k$ and During Time $D_t^k$

| $t$ | 1–12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $M_t^1$ | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 |
| $M_t^2$ | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 |
| $D_t^1$ | 0 | 1 | 2 | 1 | 3 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 |
| $D_t^2$ | 0 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 1 |

missions to be conducted. The mission types can be, for example, (1) surveillance or (2) escort for various time periods. Different mission types require different numbers of aircraft and duration.

Table IV shows the maintenance time and number of technicians of trade $r$ required to rectify failure mode $f$ for aircraft type $k$. To initiate the experiment, we specify the initial health conditions, which are represented by the estimated RUL from deep learning models, for all aircraft. Table V displays the estimated RUL (in number of operational cycles) for component $f$ of type $k$ aircraft $i$ at the beginning of the planning horizon. To take into account the maintenance time, Table VI specifies the required maintenance time for component $f$ of type $k$ aircraft at time $t$. In this example, every type of mission consumes five units of time (i.e., five operational cycles) for each component of each type of aircraft. Table VI displays the maintenance cost of repairing component/failure mode $f$ of aircraft type $k$ and the restored RUL after completing the maintenance/repair.

## B. Assessment of Model Performance

We next evaluate the scheduling performance of our proposed predictive maintenance model, compared to the traditional preventive maintenance model as a benchmark. To validate the model's performance, we conduct the tests over a 72-h horizon by assuming that the mission requirements are the same for these three days. We also assume that the RUL inputs for the second and third days are calculated based on aircraft usage and maintenance during the previous day. For example, without any maintenance, the initial RUL for the second day

Table IV
Maintenance Duration $E_{rf}^k$ and Number of Technicians $I_{rf}^k$ of Trade $r$ Required to Rectify Failure Mode $f$ for Aircraft Type $k$

| | $f=1$ | | | $f=2$ | | | $f=3$ | | |
|---|---|---|---|---|---|---|---|---|---|
| $I_{rf}^k$ | $r=1$ | $r=2$ | $r=3$ | $r=1$ | $r=2$ | $r=3$ | $r=1$ | $r=2$ | $r=3$ |
| $k=1$ | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| $k=2$ | 1 | 2 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| | $f=1$ | | | $f=2$ | | | $f=3$ | | |
| $E_{rf}^k$ | $r=1$ | $r=2$ | $r=3$ | $r=1$ | $r=2$ | $r=3$ | $r=1$ | $r=2$ | $r=3$ |
| $k=1$ | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| $k=2$ | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |

Table V
Initial Estimated RUL $r_{if}^k$ of Component $f$ for $i^{th}$ Aircraft of Type $k$

| | $i=1$ | | | $i=2$ | | | $i=3$ | | |
|---|---|---|---|---|---|---|---|---|---|
| $r_{if}^k$ | $f=1$ | $f=2$ | $f=3$ | $f=1$ | $f=2$ | $f=3$ | $f=1$ | $f=2$ | $f=3$ |
| $k=1$ | 12 | 12 | 25 | 12 | 12 | 25 | 13 | 25 | 7 |
| $k=2$ | 12 | 25 | 25 | 22 | 20 | 3 | 25 | 25 | 5 |
| | $i=4$ | | | $i=5$ | | | $i=6$ | | |
| $r_{if}^k$ | $f=1$ | $f=2$ | $f=3$ | $f=1$ | $f=2$ | $f=3$ | $f=1$ | $f=2$ | $f=3$ |
| $k=1$ | 4 | 8 | 15 | 3 | 25 | 20 | 3 | 10 | 12 |
| $k=2$ | 4 | 5 | 25 | 3 | 25 | 20 | 3 | 25 | 25 |

is the initial RUL for the first day minus the total mission duration of the first day.

Figure 6 shows the resulting optimal maintenance schedule and flight planning decisions obtained based on the mission requirements in the example scenario over a three-day planning period. In the figure, the red and green bar represent maintenance activity and mission assignment activity, respectively. The length of the bars stands for the maintenance/mission duration. The locations for maintenance activity (red bar) indicate which component is maintained. For example, for the fifth type 2 aircraft, it will maintain the first component (bottom of row 5-2) on the second day, while for the third type 1 aircraft, it will maintain the third component (top of row 3-1) on the third day. The numbers under rows MissionReq-1 and MissionReq-2 represent the number of required aircraft for each type of aircraft at a specific time slot. We can observe that all mission requirements are satisfied at all times, and there is no time conflict between maintenance and flight decisions. For example, at $t = 16$, the commander requires two type 1 and two type 2 aircraft to conduct a mission. We can see that the optimal scheduler assigns numbers 5 and 6 type 1 aircraft and numbers 5 and 6 type 2 aircraft to conduct the mission. We can interpret the decisions in terms of aircraft number 5 of type 1 as an example. It is assigned to conduct missions at time $t = 16$, and 24 in the first day. Then it is scheduled to maintain/repair component 1 at $t = 8$ in the second day, which takes 2 h to complete. After completing the repair, this aircraft is available again and can be assigned to conduct new missions. As shown in the figure, the aircraft is assigned to conduct missions at time $t = 14$, 16, 19, 20, 22, and 23.

Table VI
Maintenance Cost $c_f^k$ for Component $f$ of Type $k$ Aircraft and Restored RUL $r_{\max,f}^k$ for Component $f$ of Type $k$.

| $c_f^k$ | $f=1$ | $f=2$ | $f=3$ |
|---|---|---|---|
| $k=1$ | 100 | 90 | 85 |
| $k=2$ | 80 | 120 | 110 |
| $r_{\max,f}^k$ | $f=1$ | $f=2$ | $f=3$ |
| $k=1$ | 50 | 70 | 90 |
| $k=2$ | 50 | 80 | 100 |

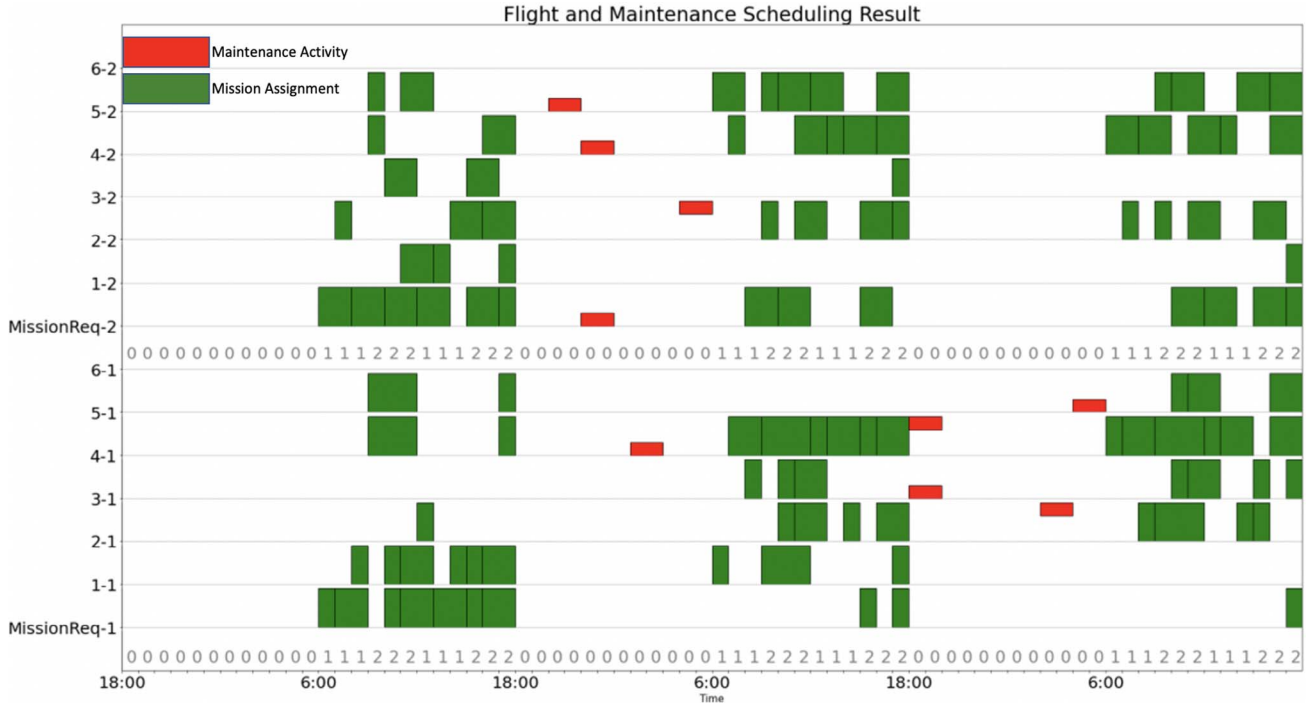



Figure 6. Flight and maintenance scheduling results.

To further demonstrate the benefits of the proposed methodology, we compare the performance of our predictive maintenance approach with that of traditional preventive maintenance methods. In preventive maintenance, an aircraft is maintained after a fixed number of operation cycles. Predictive maintenance, on the other hand, schedules the maintenance only when needed based on the predicted RUL, as described previously. In addition, in practical operations, some random failures may occur even when an aircraft is in good condition. We thus use a Bernoulli random variable, whose success probability is proportional to the cumulative usage of a component, to model these random failures. A success of the Bernoulli random variable indicates the failure of a component in our experiments. The Bernoulli random variable is realized during the preflight and postflight checks. Thus, an aircraft cannot conduct missions if a random failure occurs (or is detected) during the preflight check, and the mission is tagged as incomplete. If a random failure occurs, the aircraft needs to undergo unscheduled maintenance when maintenance resources are available; otherwise, the aircraft will be put into a waiting queue and will be unavailable to conduct any missions.

We compare the performance of predictive and preventive maintenance models over a 240-day horizon in terms of mission accomplishment rate and aircraft availability rate. The mission accomplishment rate is defined as the percentage of desired missions that are completed on time, while the aircraft availability rate is defined as the percentage of aircraft that are ready and available for mission assignments. We also evaluate the impact of the number of maintenance stations ($S$) and the duration of

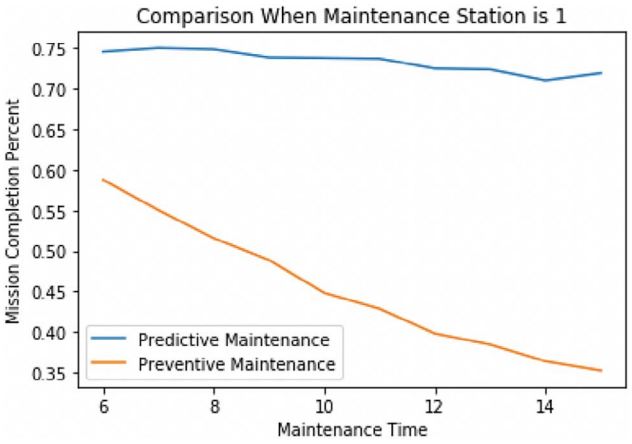



Figure 7. Mission completion rate with $S = 1$.

maintenance time ($E^k_{fr}$) on the scheduling decisions and model performance. Please note that the testing scenario involved maintenance durations ranging from 6 to 15, which differs from the specific parameters used to generate Fig. 6. This intentional difference was introduced to assess the robustness of the scheduling decisions when confronted with varying maintenance durations.

With one maintenance station ($S = 1$), Figs. 7 and 8 show the comparisons between predictive maintenance and preventive maintenance models across different maintenance durations in terms of mission accomplishment rate and aircraft availability rate, respectively. Figs. 9 and 10 illustrate the comparisons with two available maintenance stations ($S = 2$). Two important insights can be derived from the figures. First, the proposed predictive maintenance model dominates the

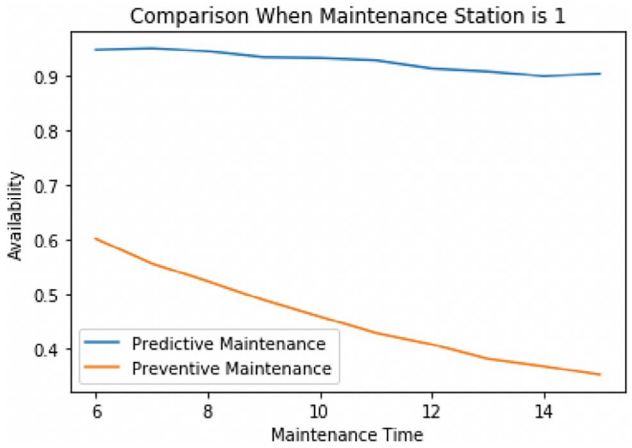



Figure 8. Aircraft availability with $S = 1$.

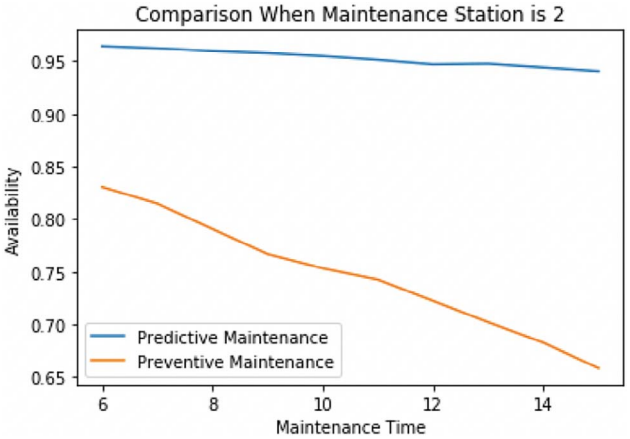



Figure 10. Aircraft availability with $S = 2$.

classical preventive maintenance model by obtaining significantly higher values in both mission accomplishment rate and aircraft availability rate. For example, when $S = 1$, the mission accomplishment rate of the predictive model ranges from 70% to 75%, while the highest value of the preventive model is only 59% and decreases dramatically as maintenance time increases. Similarly, the predictive model leads to an aircraft availability rate consistently over 90%, while the preventive model can only maintain 60% when maintenance time is short. The availability rate of the preventive model drops rapidly when maintenance duration increases. This observation leads to the second insight that the predictive maintenance model performs consistently well across different maintenance times, while the preventive model is very sensitive to required maintenance time and its performance drops significantly when maintenance time increases.

## C. Impact of Valid Inequality on Computational Efficiency

In our proposed framework and numerical experiments, the scheduling optimization problem is supposed to be implemented every 24 h. Computational efficiency thus plays an important role in the effective implementation of real-world applications. In this section, we will demonstrate the computational efficiency of the proposed solution approaches and the impact of two classes of valid inequalities.

We adopt one problem instance using the same parameter settings specified in Section V-A. We also randomly generated four other problem instances to demonstrate the performance of the proposed valid inequalities, in which the parameters $M_t^k$ and $D_t^k$ are generated according to discrete uniform distributions $U(1, 3)$ and $U(1, 3)$, respectively, to emulate small mission durations and dense mission requirements. For each problem instance, we solve it five times, aiming to avoid the possibility that the improvement is due to some random procedure within the GUROBI optimizer. We set the running time limit to 10 h (36,000 CPU seconds). We report the average solution time over the five implementations in Table VII. The first column in Table VII stands for the problem instance index. The second and third columns represent the solution time without any valid inequalities and solution time with valid inequalities proposed in Propositions 1 and 2. The results clearly indicate that the proposed valid inequalities can improve the computational speed by several orders of magnitude. For cases 1 and 5, the original formulation cannot solve the problem to exactness within 10 h. However, by incorporating the valid inequalities, the average solution time reduces to 30.0 and 39.0 s, leading to an efficiency improvement of several orders of magnitude.

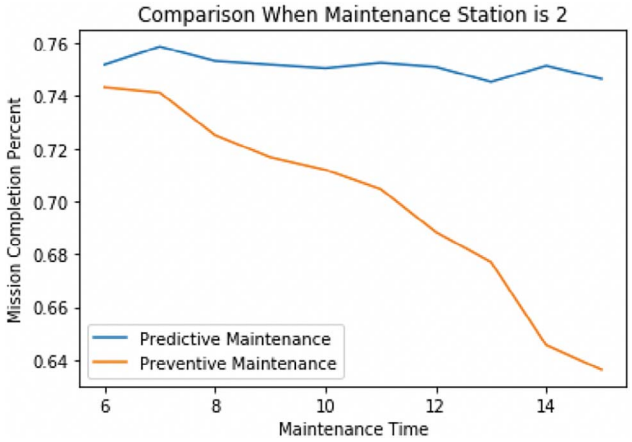



Figure 9. Mission completion rate with $S = 2$.

Table VII
Solution Time of Formulation Without and With Valid Inequalities

| Instance | Original formulation | With valid inequalities |
|---|---|---|
| 1 | 36,000 | 30.0 |
| 2 | 282.2 | 17.0 |
| 3 | 850.8 | 16.8 |
| 4 | 36,000 | 426.4 |
| 5 | 36,000 | 39.0 |

For case 4, the solution time with valid inequalities is 426.4 s, which is still approximately 80 times faster than the original formulation. For cases 2 and 3, which can be solved quickly by the original formulation within 282.2 and 850.8 s, the valid inequalities can still improve the computational speed significantly and obtain the solutions within 17.0 and 16.8 s. In sum, the two classes of valid inequalities proposed can efficiently improve the computational speed and make the proposed predictive maintenance models more easily applied in practical settings in a timely manner.

## VI. CONCLUSION

In this study, we propose an integrated learn-then-optimize framework for CBPM scheduling and flight mission planning. The bidirectional LSTM deep learning techniques are used to combine data from multiple sensors to predict the RUL values of a multicomponent aircraft. With the predicted RUL values, a MIP formulation is then proposed to maximize the fleet availability rate subject to different mission requirements and trade types. Two classes of valid inequalities are proposed to further improve the computational efficiency of the model. The proposed predictive maintenance method significantly outperforms the traditional preventive maintenance method in a hypothetical but somewhat realistic scenario.

## REFERENCES


[1] S. A. Aye and P. Heyns
"An integrated Gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission,"
*Mech. Syst. Signal Process*, vol. 84, pp. 485–498, 2017.

[2] T. P. Carvalho, F. A. Soares, R. Vita, R. d. P. Francisco, J. P. Basto, and S. G. Alcalá
"A systematic literature review of machine learning methods applied to predictive maintenance,"
*Comput. Ind. Eng.*, vol. 137, 2019, Art. no. pp. 106024.

[3] P. Cho
"Optimal scheduling of fighter aircraft maintenance."
M.S. thesis, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, USA, 2011.

[4] Z. M. Çınar, A. A. Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei
"Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0."
*Sustainability*, vol. 12, no. 19, 2020, Art. no. 8211.

[5] M. Conforti, G. Cornuéjols, and G. Zambelli
*Integer Programming*, vol. 271. Berlin, Germany: Springer, 2014.

[6] I. de Pater and M. Mitici
"Predictive maintenance for multi-component systems of repairables with remaining-useful-life prognostics and a limited stock of spare components,"
*Rel. Eng. Syst. Saf.*, vol. 214, 2021, Art. no. 107761.

[7] I. de Pater, A. Reijns, and M. Mitici
"Alarm-based predictive maintenance scheduling for aircraft engines with imperfect remaining useful life prognostics,"
*Rel. Eng. Syst. Saf.*, vol. 221, 2022, Art. no. 108341.

[8] Q. Feng, X. Bi, X. Zhao, Y. Chen, and B. Sun
"Heuristic hybrid game approach for fleet condition-based maintenace planning,"
*Rel. Eng. Syst. Saf.*, vol. 157, pp. 166–176, 2017.

[9] A. Gavranis and G. Kozanidis
"An exact solution algorithm for maximizing the fleet availability of a unit of aircraft subject to flight and maintenance requirements,"
*Eur. J. Oper. Res.*, vol. 242, pp. 631–643, 2015.

[10] A. Gavranis and G. Kozanidis
"Mixed integer bi-objective quadratic programming for maximum-value minimum-variability fleet availability of a unit of mission aircraft,"
*Comput. Ind. Eng.*, vol. 10, pp. 13–29, 2017.

[11] F. Gers, J. Schmidhuber, and F. Cummins
"Learning to forget: Continual prediction with LSTM,"
*Neural Comput.*, vol. 12, pp. 2451–2471, 2000.

[12] L. Guo, N. Li, F. Jia, Y. Lei, and J. Lin
"A recurrent neural network based health indicator for remaining useful life prediction of bearings,"
*Neurocomputing*, vol. 240, pp. 98–109, 2017.

[13] S. Hochreiter and J. Schmidhuber
"Long short-term memory,"
*Neural Comput.*, vol. 9, pp. 1735–1780, 1997.

[14] G. Kozanidis
"A multiobjective model for maximizing fleet availability under the presence of filight and maintenance requirement,"
*J. Adv. Transp.*, vol. 43, pp. 155–182, 2009.

[15] G. Kozanidis, A. Gavranis, and G. Liberopoulos
"Heurisitics for flight and maintenance planning of mission aircraft,"
*Ann. Operations Res.*, vol. 221, pp. 211–238, 2014.

[16] G. Kozanidis, G. Liberopoulos, and C. Pitsilkas
"Flight and maintenance planning of military aircraft for maximum fleet availability,"
*Mil. Operations Res.*, vol. 15, pp. 53–73, 2010.

[17] P. Kundu, A. K. Darpe, and M. S. Kulkarni
"An ensemble decision tree methodology for remaining useful life prediction of spur gears under natural pitting progression,"
*Struct. Health Monit.*, vol. 19, no. 3, pp. 854–872, 2020.

[18] J. Lee and M. Mitici
"Multi-objective design of aircraft maintenance using Gaussian process learning and adaptive sampling,"
*Rel. Eng. Syst. Saf.*, vol. 218, 2022, Art. no. 108123.

[19] Z. Li, J. Guo, and R. Zhou
"Maintenance scheduling optimization based on reliability and prognostics information,"
*Annu. Rel. Maintainability Symp.*, 2016, pp. 15.

[20] K. T. Nguyen and K. Medjaher
"A new dynamic predictive maintenance framework using deep learning for failure prognostics,"
*Rel. Eng. Syst. Saf.*, vol. 188, pp. 251–262, 2019.

[21] F. Peschiera, O. Battaia, A. Hait, and N. Dupin
"Long term planning of military aircraft flight and maintenance operations," 2020, arXiv: 2001.09856.

[22] F. Peschiera, R. Dell, J. Royset, A. Hait, N. Dupin, and O. Battaia
"A novel solution approach with ML-based pseudo-cuts for the flight and maintenance planning problem,"
*OR Spectr.*, vol. 43, 2021, pp. 635–664.

[23] R. R. Richardson, M. A. Osborne, and D. A. Howey
"Gaussian process regression for forecasting battery state of health,"
*J. Power Sources*, vol. 357, pp. 209–219, 2017.

[24] N. Safaei, D. Nanjevic, and A. K. Jardine
"Workforce-constrained maintenance scheduling for military aircraft fleet: A case study,"
*Ann. Operations Res.*, vol. 186, pp. 295–316, 2011.

[25] A. Saxena, K. Goebel, D. Simon, and N. Eklund
"Damage propagation modeling for aircraft engine run-to-failure simulation,"
*Proc. IEEE Int. Conf. Prognostics Health Mangement*, 2008, pp. 1–9.

[26] M. Schuster and K. Paliwal
"Bidirectional recurrent neural networks,"
*IEEE Trans. Signal Proess.*, vol. 45, no. 11, pp. 2674–2681, Nov. 1997.

[27] O. Serradilla, E. Zugasti, J. Rodriguez, and U. Zurutuza
"Deep learning models for predictive maintenance: A survey, comparison, challenges and prospects," Appl. Intell., vol. 52, pp. 1093410964, 2022

[28] C. Shen, D. Wang, F. Kong, and W. T. Peter
"Fault diagnosis of rotating machinery based on the statistical parameters of wavelet packet paving and a generic support vector regressive classifier,"
*Measurement*, vol. 46, no. 4, pp. 1551–1564, 2013.

[29] S. Varsamopoulos, K. Bertels, and C. G. Almudever
"Comparing neural network based decoders for the surface code,"
*IEEE Transactions on Computers*, vol. 69, no. 2, pp. 300–311, 2019.

[30] H. Wang, X. Ma, and Y. Zhao
"An improved Wiener process model with adaptive drift and diffusion for online remaining useful life prediction,"
*Mech. Syst. Signal Process*, vol. 127, pp. 370–387, 2019.

[31] X. Wang
"Wiener processes with random effects for degradation data,"
*J. Multivariate Anal.*, vol. 101, no. 2, pp. 340–351, 2010.

[32] Y. Wang, S. Limmer, D. Van Nguyen, M. Olhofer, T. Bäck, and M. Emmerich
"Optimizing the maintenance schedule for a vehicle fleet: A simulation-based case study,"
*Eng. Optim.*, vol. 54, no. 7, pp. 1258–1271, 2022.

[33] Y. Wen, M. F. Rahman, H. Xu, and T.-L. B. Tseng
"Recent advances and trends of predictive maintenance from data-driven machine prognostics perspective,"
*Measurement*, vol. 187, 2022, Art. no. 110276.

[34] B. Yang, R. Liu, and E. Zio
"Remaining useful life prediction based on a double-convolutional neural network architecture,"
*IEEE Trans. Ind. Electron.*, vol. 66, no. 12, 9521–9530, Dec. 2019.

[35] Z.-S. Ye, Y. Shen, and M. Xie
"Degradation-based burn-in with preventive maintenance,"
*Eur. J. Oper. Res.*, vol. 221, no. 2, 360–367, 2012.

[36] J. Zhang, P. Wang, R. Yan, and R. X. Gao
"Deep learning for improved system remaining life prediction,"
*Procedia CIRP*, vol. 72, pp. 1033–1038, 2018.

[37] J. Zhang, P. Wang, R. Yan, and R. X. Gao
"Long short-term memory for machine remaining life prediction,"
*J. Manuf. Syst.*, vol. 48, pp. 78–86, 2018.

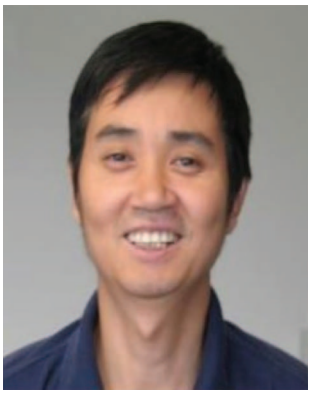
[38] Y. Zhang, R. Xiong, H. He, and M. G. Pecht
"Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries,"
*IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5695–5705, Jul. 2018.

[39] T. Zonta, C. André da Costa, R. da Rosa Righi, M. José de Lima, E. Silveira da Trindade, and G. P. Li
"Predictive maintenance in the industry 4.0: A systematic literature review,"
*Comput. Ind. Eng.*, vol. 150, 2020, Art. no. 106889.

**Zhengyang Fan** received the M.S. degree in statistical science from George Mason University, Fairfax, VA, USA, in 2017 and he is currently a Ph.D. candidate in the Department of Systems Engineering and Operations, at George Mason University, Fairfax, VA, USA. His current research interests focus on explainable AI, stochastic programming, and distributionally robust optimization.

**Kuo-Chu Chang** received the M.S. and Ph.D. degrees in electrical engineering from the University of Connecticut, Storrs, CT, USA, in 1983 and 1986, respectively. He is currently a Professor with the Department of Systems Engineering and Operations Research, George Mason University, Fairfax, VA, USA. His research interests include estimation theory, data fusion, Bayesian inference, machine learning, and financial engineering. He has more than 35 years of industrial and academic experience and has published more than 300 papers in the areas of multitarget tracking, distributed sensor fusion, Bayesian networks technologies, and quantitative finance. As an IEEE Fellow, his research has been strongly supported by ONR, AFOSR, NSF, DARPA, and ARO.

**Ran Ji** is an Assistant Professor with the Department of System Engineering and Operations Research, George Mason University, Fairfax, VA, USA. His research focuses on stochastic optimization, distributionally robust optimization, chance-constrained programming and uncertainty quantification, and data-driven decision-making. Specifically, some of his research goals include tackling optimization problems under uncertainty via mathematical programming. His application areas of interest include financial risk modeling, portfolio optimization, supply chain risk analytics, inventory and production planning, statistical learning, and data mining.



**Genshe Chen** received the B.S. and M.S. degrees in electrical engineering and the Ph.D. degree in aerospace engineering, in 1989, 1991, and 1994, respectively, all from Northwestern Polytechnical University, Xian, China. He is currently the Chief Technological Officer (CTO) of Intelligent Fusion Technology, Inc., Germantown, MD, USA, where he directs research and development activities for the Government Services and Commercial Solutions group. He was the CTO of DCM Research Resources LLC, Germantown, MD, USA, and the program manager for networks, systems, and control at Intelligent Automation, Inc. His research interests include space resiliency, space situational awareness, cognitive radio for satellite communications, cooperative control and optimization for military operations, multi-INT data fusion, guidance, navigation, and control, electronic warfare, game theory, and human-cyber-physical systems.