Journal of Industrial and Management Optimization

JIMO, 22(9): 4481–4513.
DOI:10.3934/jimo.2026156
Received: 30 June 2026
Revised: 08 August 2026
Accepted: 17 August 2026
Published: 24 August 2026

https://www.aimspress.com/journal/jimo

*Research article*

# A reinforcement learning-based rolling optimization method for weekly aircraft maintenance scheduling under repair efficiency uncertainty

**Shihong Li[1,2], Pengliang Che[1], Mengjin Qu[1] and Qing Li[1,*]**

[1] Department of Automation, Tsinghua University, Beijing 100083, China
[2] Aviation University of Air Force, Changchun 130022, China

* **Correspondence:** Email: liqing@tsinghua.edu.cn.

## Appendix A. Executable synthetic-data generation and replay protocol

A1. Initialize a fresh AnyLogic model instance for the specified scenario, fleet size, method, and seed. Each evaluation cell, therefore, begins from an independent simulator state.

A2. For each week, generate/replay the exogenous task demand before action selection. In the base generator, an approximately unit-mean Gaussian multiplier is clamped to [0.70,1.50], multiplied by 180, and rounded to an integer, and the resulting demand is clamped to [60,300]. The exported week-level demand is the reproducibility record.

A3. Generate/replay eta before action selection. The main experiment uses the truncated-normal model with the specified sigma and support [0.70,1.50]; distribution-family tests substitute the bounded uniform or support-truncated lognormal model defined in Section 5.5.

A4. Compute the current observation and executable capacity. For the full-state PPO, the current eta is observable; for the no-efficiency-state policy, the policy feature is neutralized, but the real eta remains active in the executor.

A5. Generate the method action from the current information set. FCFS and Greedy apply their deterministic rules; the exact MILP and GA solve/search the predefined current-week objective; PPO performs fixed-checkpoint inference. The current action uses current demand and eta, with future realizations outside the decision state.

A6. Apply common physical execution semantics, including operational state checks, eta/margin capacity limits, legal maintenance transitions, nonnegative/integer execution, and the at-most-one proactive maintenance rule.

A7. Update backlog, aircraft/maintenance states, availability, cost, constraint pressure indicators, and the next-week state; export the week-level raw record.

A8. Repeat the weekly cycle for 50 weeks, terminate the process/JVM/model cell, and compute seed-level summaries. Methods compared within the same experiment share the same exogenous demand and eta records. Retained sequence files, weekly records, seed summaries, input-matching records, and process lifecycle logs provide the reproducibility trail for the reported comparisons.




AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)