

Review: Reinforcement-Learning-Based Collaborative Optimization of Load Regulation and Opportunistic Maintenance of Wind Turbines (wes-2026-157)

Summary

The authors propose an RL method that jointly decides daily power derating (DR) and multi-component maintenance for a wind turbine over a 19-year horizon. The policy is trained with a PPO variant in a simulator built from synthetic wind, a normalized-power load proxy, and Weibull reliability curves.

I like the idea. Letting operational control slow degradation before maintenance is a sensible direction and fits WES well. The paper is also well organized. My main concern is that the results rest entirely on a simulator whose key assumption is not validated, and some reported numbers do not add up.

1. Scientific Quality: Fair

- The whole benefit of derating comes from Eq. (10), which assumes aging grows linearly with normalized power. The coefficients κ_i , α_i , μ_i are not reported or validated. In reality, fatigue depends nonlinearly on load ranges, thrust peaks near rated speed, and each component responds to different loads (torque, thrust, pitch activity).
- I am not convinced the evaluation supports the claimed gains. Training uses a single seed, all strategies are tested on the same wind sequence, and the final policy is selected on that same setting. A 2% profit gap may simply be noise. The baselines also use fixed, untuned thresholds, and there is no rule-based DR baseline, so I cannot tell whether the gains come from derating or from RL. Please report results over several seeds and held-out wind realizations, and add a simple CBM-plus-DR baseline.
- I tried to check the cost numbers and could not make them work. Table 2 says a minor maintenance costs 0.21 M CNY and a major one costs 1.6 M CNY. RL-OppOM performs 29 minor and 8 major actions, which already adds up to about 18.9 M CNY. Table 4 reports only 5.93 M CNY. The same mismatch appears for CBM. I may be missing something, but the authors should show exactly how the costs in Table 4 are computed. A similar issue appears for OppM's availability, which I could not reproduce from its downtime.
- Computational inconsistencies like these, in a study built entirely on simulated data, make a **strong case for publishing the code**.

2. Presentation Quality: Fair

- The abstract sentences are long and hard to follow.

- Figure 1 is blurry in places and uses symbols defined only later. A simpler schematic would work better.
- Section 2.1: why simulate wind when SCADA data are available? I assume a 19-year series is needed, but please say so and explain how ρ , σ , v_0 , and the seasonal factors were calibrated. Eq. (2) is a Gaussian AR(1) process, so it is unclear how it yields a Weibull marginal.
- Figure 2 looks unusually clean for turbine SCADA data. Please overlay the measured histogram.
- Please explain where the values in Tables 1, 2, and 3 come from. The Table 1 Weibull parameters are from an offshore study, so their use for an onshore turbine in northern China needs justification.
- Several parameters are missing (R_{ref} , R_{near} , R_{new} , R_{low} , v_{gust} , F_{th} , penalty coefficients, score weights).
- The methods are written mostly in past tense. Present tense reads more naturally for model definitions.
- Small fixes:
 - Eq. (4) should end at v_{out} , not v_{in} , and Eq. (6) is garbled.
 - L. 166 should read $L(d)$, not $F(d)$.
 - η , λ , and β are each reused with two meanings.
 - Eq. (17) contains a weird character, and Eq. (32) shows “ $\div 2$ ”.
 - Equation formatting (Eqs. 3, 7, 13, 17) is inconsistent.
 - The Fig. 5 legend reads “OppO M”.