This manuscript reviews studies on the application of Deep Reinforcement Learning (DRL) for offshore wind-farm maintenance planning and highlights the potential benefits of DRL compared to traditional maintenance strategies. The topic of implementing data-driven approaches for operation and maintenance is highly relevant for the cost-effective deployment of offshore wind farms. Also, the overall structure of the manuscript is organized, with the abstract providing an informative summary of the content.

As stated in Line 474, "real-world applications of DRL in offshore wind O&M are still in early stages", most of the studies reviewed rely solely on simulations. To provide a more balanced perspective, the manuscript should not only highlight the potential of DRL, but also acknowledges the limitations, assumptions, and simplifications inherent in these simulation-based approaches. Also, some discussions lack sufficient evidence and requires further clarification. The following comments are offered with the aim of improving the quality of the manuscript.

1. Although this paper is a review rather than a simulation study that typically states the research gap in previous studies, it would benefit from certain discussions on how previous reviews or studies have assessed DRL, RL, or machine-learning methods for wind-farm maintenance and operations. A brief summary of prior reviews would be sufficient to strengthen the novelty of this paper and help to clarify its unique contribution.

2. The internal hyperlinks, which are clickable texts that jump to a reference or figure, are not required, but the manuscript would benefit from properly functioning clickable cross-references. Now in the manuscript, the in-text references to figures, tables, or citations are not linked. While a minor issue, this requires the readers manually locate the referenced material. Also, the reference list duplicated entries (e.g., Lines 585–590), which should be removed.

3. In both abstract and introduction sections, the authors mentioned the limitations of using traditional approaches for offshore wind-farm maintenance, and emphasize the benefits of using DRL over both classic RL and traditional methods.
   However, the manuscript does not discuss the potential disadvantages or limitations of DRL compared to these remaining approaches. For a neutral discussion, the authors should also briefly address the general challenges associated with DRL for wind-farm maintenance, such as its black-box nature and computational complexity.
   Additional suggestion is, the manuscript should discuss the potential difficulties of applying DRL in real-world offshore wind-farm maintenance planning. Challenges such as high computational requirements would affect practical deployment. Including these considerations would provide a more objective review.

4. The authors mention (Lines 50–51) that they reviewed single-agent DRL settings to highlight the performance of key algorithms in discrete decision-making tasks. However, in Lines 75–80, they state that policy-gradient methods improve stability and sample efficiency in large or continuous action spaces. It is unclear how this advantage applies to discrete decision-making: are policy-gradient methods also expected to outperform or normally perform in discrete settings, or is their advantage primarily limited to continuous or high-dimensional action spaces? A clarification on the relevance of policy-gradient performance in discrete tasks would help to understand these statements.

5. How should Figure 1 be interpreted? The first row represents the challenges, while the second row shows the opportunities. However, it is unclear why the arrows point from the challenges to the opportunities—shouldn't the opportunities be presented as ways to address these challenges? Similarly, in Figure 7, should the listed items—such as knowledge of aerodynamics—be considered as contributors to offshore operation and maintenance (O&M)?

   Also, the authors suggest that proactive scheduling cannot handle uncertainty from harsh environmental conditions, but this is a bit questionable. For example, proactive scheduling informed by operational data, such as lidar- measured wind data, could provide valuable insights and help to lower the influence of environmental uncertainty.

   Similarly, it is unclear why the adaptive decision-making and learning have no impact on complex logistic resource allocation (as no line to connect them in Figure 1). For instance, adaptive methods that update decisions based on real-time information—such as vessel availability, crew schedules, should enhance resource utilization. A clarification on these points would help to understand better about the challenges and opportunities using DRL in offshore wind sector.

   Finally, as the focus of the paper is on 'maintenance planning', it is unclear whether the discussion is intended to concentrate solely on 'proactive scheduling', or whether it aims to address a broader combination of these opportunities outlined in Figure 1(e.g., adaptive decision-making, learning-based optimization, condition-driven strategies, or resource coordination). In case the maintenance planning is framed narrowly as 'proactive scheduling', or broadly as a combined set of four opportunities, would help readers to better understand the scope and contribution of the review.

6. The overall structure of the manuscript is good, but the content in Lines 135–140 could be better organized. It may be more effective or proper to place this content in a separate subsection or to just include an additional table comparing the three single-agent DRL approaches. These would improve clarity and allow readers to quickly understand the differences or limitations of each method.

7. Lines 141 has a small typographical error: there is insufficient spacing between "experiences,even."

8. In Line 149, the phrase "As our environment grows in size" is understandable, but it could be improved for clarity. For example, it could be rephrased as "As the environment increases in size as wind farms scale up …" to make it more precise for a scientific context.

9. In Lines 312–313, the phrase "noticeable reduction in lifecycle cost" is too vague. It is unclear how much the cost is reduced. A quantitative estimate, similar to the example given in Line 315, would make this statement more precise for the reader.

10. In Lines 449–451, the authors state: "By integrating these domain specifics, summarized in Figure 7, DRL models become more realistic and effective." It is unclear why incorporating more considerations necessarily makes the DRL more effective. The examples or referenced studies in the

manuscript only illustrate the specific considerations, but no quantitative evidence of improved efficiency is provided.
Moreover, involving these additional considerations may increase training effort and model complexity, which is a critical concern for DRL applications. Regarding the statement: “Guiding the learning with known physics (wake loss), operational constraints (weather, logistics), and expert knowledge (PHM interpretations) greatly improves performance”, it is worth questioning whether over-modeling could potentially reduce performance. For example, conflicts between information sources or increased computational burden might make the model harder to train or even degrade its effectiveness.

11. In Lines 470–480, the authors cite a reference study to support the claim that DRL is applicable to real-world wind-farm maintenance, as it can improve power performance through dedicated maintenance planning. However, it is important to note that the study used the Jensen wake model, which cannot accurately capture wake dynamics and therefore may not reliably predict power output, especially for close-spaced turbines due to its limitations in the near-wake region. Since real-world applications of DRL in wind-farm maintenance are still at an early stage, the limitations of the current models and assumptions should be presented to provide a more neutral assessment.

12. A suggestion regarding the presentation of the three main families of DRL algorithms in Lines 65–80: it would be clearer and more readable if they were listed in a structured format, similar to how the potential avenues are presented in Lines 500–515. Using a concise, bullet-point or tabular style would help readers quickly grasp the distinctions between the algorithm families.