

Response to reviewers

Reviewer 1

The paper proposes the development of an open innovation ecosystem, WeDoWind, to foster open science and open innovation in wind energy. This is done by transforming the WeDoWind from a platform for documenting and managing challenges to an open innovation ecosystem using the Design Thinking approach. The viability of the ecosystem is determined based on a case study. Then, feasibility study KPIs and case study KPIs are defined and evaluated. The proposed ecosystem looks interesting. Nevertheless, some improvements could be made to clarify the paper.

Thank you for taking the time to carry out this review! Modifications based on these comments are given in red in the new version of the paper.

The definition of KPIs is sometimes abstract and subjective. This is true for both types of KPIs. For example, the Network Maturity KPI, which is measured by the number and type of interactions between participants, obviously depends on the number of participants. If it were defined per participant, it could be much clearer. The same could be true for the Community Traction KPI. It is therefore recommended that KPI definitions be reviewed and ways found to express them in a more objective and assessable manner.

To address all the comments on the KPIs, we decided to replace KPIs with the more traditional TELOS aspects and undertake a more qualitative analysis. The entire evaluation sections have been rewritten.

Assessing viability through a case study may be appropriate. However, the case study description and its results distract from the article's ultimate objective. It would be advisable to focus that section more on how the case study was developed, managed, and evaluated within the ecosystem rather than on its description.

Details of the case study description and the solutions have been removed or shortened.

When evaluating KPIs, it is important to avoid subjective criteria (again) as much as possible. For example, when evaluating the KPI for comparison metrics, a 'Low-Medium' rate is assigned when it is clear that the metrics cannot be directly compared because participants provided them in heterogeneous formats, and one participant did not provide them at all. Therefore, this KPI should be rated as 'Low'.

See comment about KPIs above

A clearer explanation of some KPIs is needed. For example, in the same KPI as in the previous point (Comparative Metrics), the fact that participants provided such heterogeneous metrics may be due to a flaw in the case definition.

See comment about KPIs above

Reviewer 2

Its great to see a paper that not only encourages open innovation, but that demonstrates this is possible through a case study. It is also nice to see this in practice in the paper, with both code and data also being made accessible.

Thank you for taking the time to carry out this review! Modifications based on these comments are given in blue in the new version of the paper.

I think some improvements could be made. For example it is unclear if the case study is being assessed here against the ability to foster open science within WeDoWind, or if the results of the case study are being assessed directly (i.e. how well the solutions work). While the later is a good success case demonstrating that the process works, I feel this paper should focus more on the ability of the case study example to be run in WeDoWind and foster open science, and the lessons learnt in doing so, e.g. what type of KPIs are useful, how the community can be engaged, how a challenge is best run, the usability of results, demonstration of market demand, etc., and other feedback into the feasibility KPIs and ecosystem.

We have removed some details of the case study description as well as the submitted solution, in order to make this more clear.

The actual solutions and results of the case study could then be a paper in their own right, and this paper could be more focused, so perhaps it is worth splitting the paper in two?

Agreed, and we have removed some details of the submitted solutions.

The conclusion could also be strengthened by highlighting the lessons learnt during the case study, and also by referring back to the KPIs that the WeDoWind platform was assessed against, so that the knowledge learnt through this work is applied.

We have expanded the conclusions and referred to the evaluation aspects (previously KPIs – altered due to comments of the other reviewer). We added some “lessons learned” to the conclusions as well.

Line 38: slightly odd combo, as one is an entity and the other a process, perhaps "companies & universities"?

Changed

Line 113: typo: included

Changed

Line 133: Many of the KPIs are subjective and/or cover multiple measures. This may make it difficult to assess and compare the results objectively, and to repeat the

research and make fair comparisons (e.g. with a follow up study). There are also a large number of KPIs, which might make it difficult to track. The targets, i.e. what is a good/bad result, are also not stated here, so the review of each aspect feels subjective.

We have improved the whole KPI analysis section, also because of the comments of the other reviewer. Instead of KPIs, we use “evaluation aspects” and focus the TELOS study on a qualitative analysis of the strengths and weaknesses.

Line 168: This sentence might need re-wording.

Added the word “and”

Line 185: This sentence may also need re-wording

Added a colon to make more clear

Line 247: Should this section be split into a "Method descriptions" and "KPI Comparison Metrics" section? I'm also unsure if the detailed method descriptions are required, as they feel both too detailed and too brief, and so perhaps detracts from the focus on the case study KPIs. This section could then better compare and analyse the metrics.

We removed the detailed method descriptions

Line 314: The first paragraph compares the Advantages and Disadvantages of each approach individually, but doesn't show clearly how the Advantages and Disadvantages can be compared across the methods, and so the benefits of this KPI

This is exactly what Table 2 is showing

Line 391: Why is this KPI rated "Medium" when it not only provided insights, but also clear comparison to other challenges?

The KPI analysis has been changed to a qualitative assessment, as mentioned above

Line 401: But this KPI helps clearly track, with quantitative values, New Activity, so I'm uncertain why the KPI ranked low. Is it because this KPI isn't being used to drive a decision process?

The KPI analysis has been changed to a qualitative assessment, as mentioned above

Table 1: It feels like the KPI Comparison Metrics / KPI Advantages and Disadvantages were inconsistent, as not all methodologies were compared on the same metric.

Agreed, and this is why we write “It can be seen in Table 1 that the comparison metrics (accuracy, precision, recall and F-scores) could be used to some extent to compare the performance of the different methods. The fact that some reports did not include these metrics specifically, or they were given for multiple different cases, made it more difficult to make a direct comparison.” when discussing the KPI Comparison Metrics.

Table 3: I'm not certain if the value of each KPI is being assessed, or whether this is showing the output of each KPI for the case study. For example Participant Satisfaction and New Activity appear to rate the challenge against these KPIs (e.g. the level of participant satisfaction), rather than the value of the KPIs themselves (e.g. the ability to track and so work to improve participant satisfaction). While the others appear to rate the KPI themselves, e.g. the ability of the selected Evaluation Metrics to compare results, rather than stating that the KPI results (e.g. very high accuracy) were good, so would otherwise be rated "High".

This table has been changed now due to the above-mentioned changes

Line 406: The rating against some these KPIs looks impressive, but it can be hard to interpret without a goal or baseline. For example 25000 impressions sounds impressive, the number and range of groups similarly, so what would a good level be? I understand this is difficult to define, but without context its hard to understand the rating of the KPIs, and could lead to different ratings if assessed by different authors, e.g. if this work is repeated in the future either by WeDoWind or another group.

We have improved the whole KPI analysis section, also because of the comments of the other reviewer. Instead of KPIs, we use "evaluation aspects" and focus the TELOS study on a qualitative analysis of the strengths and weaknesses.

Line 510: I think future work should relate back to the strengths and weaknesses identified in this work, as that would better demonstrate the value of this research and paper.

This has been included now!