General Comments

This work presents a new approach (NSH) to solving the wind farm layout optimization problem using a MILP approach that is made more tractable by a simplified wind farm AEP model. The results of the model and the optimization algorithm are clearly compared to previous works and seem reasonably reproduceable. The work appears to be well-founded from a scientific perspective, is relevant to the subject matter of Wind Energy Science, and provides meaningful contributions to field.

While the work is reasonably well presented, the english grammar and usage in the work present a barrier to understanding. The manuscript should be carefully, preferably professionally, edited to address these concerns so the material will be more accessible, clear, and useful to the community.

Specific Comments

Abstract

- Line 4-5: "deficit is aimed" I don't know what is meant by this.
- Line 5-6: it is unclear if the heuristic wraps the model (formulations?) or is separate. Consider clarifying.
- Line 8: This sentence was confusing to me, but I think I understand. Consider re-working. I think the intended meaning is that the results of the benchmark problems show that using some substitute objective rather than actual AEP can be a good approach.
- Line 10: "match" is probably a bit strong for the presented results, maybe say the results are competitive or something that does not indicate equality

1: Introduction

- Line 17: I don't think I'm convinced about the importance of wind farm layout optimization by this paragraph. You state that wind energy is important politically, is presumably profitable without subsidies, and is a mature industry. The profit and maturity seem to hurt the argument for why this study is important. It sounds like things are just fine without WFLO. I'd suggest re-working this first paragraph. You could consider discussing the tight margins of wind developers and OEMs, especially offshore. You could also mention some hard values for how improved wind farm layouts could reduce the cost of energy even further. Basically, be careful to lay a clear foundation for why this work matters. You don't need to cover a lot of detail or history, but do make a clear case.
- Line 24: I think you are citing Deb (2013) here for an example of a GA, but it reads like you are pointing readers to the GA that Mosetti used, only the dates don't line up (2013 vs 1994). Consider reworking this or putting the expected citation (or no citation, you already cited Mosetti

which presumably has the information on the GA)

- Line 26: Consider removing "and the associated numerical algorithms" because you are stating "main components". Nearly all computational methods will have "associated" algorithms. However, I'd argue that the wake combination model qualifies as a "main" component as well.
- Line 30-34: while the wake model background may not need to be complete, the background given here is not quite correct.
    1. I think Niayifar and Porte-Agel (2015) is mostly focused on the wake combination and turbulence intensity to extend the Bastankhah model to multiple turbines. In this light the citation would be better placed with the wake combination citations. Also note that there is a journal paper by the same authors from 2016 on the topic that may be a better source to cite here.
    2. The Jensen cosine model was actually proposed by Jensen in 1983, so it may be good to cite that paper for the Jensen cosine model as the original source, though the Thomas et al. paper does provide some clarifications.
    3. The list as given seems to show several smooth and differentiable wake models, but the combined citations seem to really only refer to two distinct wake models. I'd suggest making this a little more clear in the discussion.
    4. While the sum of squares or linear combination statement is correct to my knowledge, it may be worth mentioning that the two methods have been used with local and freestream velocity conditions. This makes for four distinct proposed wake combination methods.
        - Linear/freestream: Lissaman 1979
        - Sum of squares/freestream: Katic et al. 1986
        - Linear/local: Niayifar and Porte Agel 2015, 2016
        - Sum of squares/local: Voutsinas 1990 Update: I saw you do discuss this nuance later. It may ok as is, but it did seem incomplete to me at first.
- Line 49: the jump from gradient-based and gradient-free algorithms to discrete algorithms was not clear and needs motivation. Consider stating the connection and purpose of the jump for those unfamiliar with the algorithms ((1) discrete methods are generally a sub-set of gradient-free methods, and (2) why are we talking about them here?)
- Line 59-69: consider also citing https://wes.copernicus.org/articles/7/1137/2022/
- Line 71: how is modeling economic metrics an advantage of the discrete model? This can and has been done in a continuous space for optimization. - see https://onlinelibrary.wiley.com/doi/epdf/10.1002/we.2310
- Line 72: (ii) can be done continuously, but it is more difficult
- Line 72-73: why is (iii) specific to a discrete formulation?
- Line 73-74: what is the distinction between "cost functions" (iv) and "economic metrics" (i)?
- Line 74: fully continuous WFLO has been done with multiple turbine types
    - https://www.wind-energ-sci.net/4/99/2019/wes-4-99-2019.pdf

- Line 76-77: consider elaborating on this idea and why discrete optimization is well suited to overcome the convexity problem

2: Physic Modelling

- Line 97-99: I'm not sure how this statement "No particular ..." relates to the first sentence in the paragraph. Also, Thomas et al. 2022b specify some restrictions on the mathematical structure for controlling wake diameter and deficit, at least for their purpose. Specifically, the wake deficit and wake diameter must be separately controlled.
- Line 101: from which source? there are two references.
- Line 105: why is Thomas and Ning 2018 cited here and at line 32 for the simplified Gaussian? For the original Bastankhah wake model, I'd suggest citing Bastankhah 2016. For the simplified model, use the citation given in the following sentence (IEA Wind Task 37 2019)
- Line 110: $d_{ij}^{\parallel}$ and $d_{ij}^{\perp}$ are not used in Eq. (1), though the coordinate frame clarifications are helpful, the symbols used seem extraneous at this point in the paper. You could possibly include these symbols as additional equations following the equation explanation of Eqs. (1) and (2) in preparation for use later in the paper.
- Line 115: consider removing one of these duplicate mathematical statements.
- Line 125-130: The references used to arrive at Eqs. (6) and (7) were given in the introduction, but I think it would be helpful to provide them again here.
- Line 133: why is the power curve non-differentiable specifically at rated wind speed? The definition provided in this manuscript is non-differentiable at the rated power, but the continuity of the power curve is just dependent on the power curve definition, so this statement is not correct in general.

3: Optimization models

- Eqs. (12) to (14)
  - The presentation here is difficult to follow. Perhaps consider breaking them up into more equations with more explanation and grouping by interval (1, 2, a+1, m+1, m+2).
  - the statement "for a = 1,...,m" should be applied to each numbered equation it applies to individually.
  - I'm not sure how the delta u is supposed to be applied in Eq. (13). Re-working the presentation of these equations should help.
  - are "a" and "l" being used for the same thing here? If so, correct. If not, please clarify.
  - There may be a better way to present the interval values, the above are just my ideas at the moment.
- Eq. (16a): xi, eta, and u are specified as design variables, but I think eta and u are state variables dependent on xi, so it seems that xi represents

all design variables. I've only seen design variables represented in the sub-scripted variables under the "maximize" in the optimization equation.
- Section 3.2: this approach appears similar to the FLOWERS model found in https://wes.copernicus.org/articles/7/1137/2022/wes-7-1137-2022.pdf. I'd suggest contrasting the method in the submitted manuscript to the FLOWERS model, perhaps in the introduction, but referring back to it again here.
- Line 204: I need a little clarification regarding which "outlook" the IEA 37 studies follow.
- Line 211: please provide justification for dropping the square roots. Why is the model expected to be correct if the square root is simply "dropped"?
- Eq. 20: how did you get to `b_{i,l} + b_{l,i}` and the `l>i`? I don't see offhand how those terms come from combining eqs. 18 and 19 as stated.

4: Neighborhood search heuristic

- Alg. 1, Line 13: check spacing
- Line 263-264: what is meant by "stopped until"?
- The NSH algorithm seems similar to the one developed by Paul Malisani and presented in "A Comparison of Eight Optimization Methods Applied to a Wind Farm Layout Optimization Problem" by Thomas et al. (https://wes.copernicus.org/preprints/wes-2022-90/). Consider comparing and contrasting the approaches.

5: Computational experiments

- Line 286: why these parameter values?
- Line 299: it would be nice to see all non-default parameters (the introduction "for example" seems to indicate that only some of the non-default parameters are given). Consider putting in a table with the non-default parameter values.
- Line 305-316: was this sampling method compared to any other methods?
- Line 334: how do we know it is "still strong enough"? What was the bar?
- Line 346: which model is "exact"? All the models presented in this paper appear to be approximations.
- Line 349: Perhaps the "deterioration" is partly due to "dropping" the square root?
- Line 349: "this" is unclear, state meaning explicitly
- Table 1, Fig. 4: beautiful use and presentation of correlation. Nice work!
- Line 363: It would be helpful to provide more information about the tuning process.
- Line 363: My understanding of C, T, and V was incomplete and I had to go back and re-read previous sections and this sections to get straitened out. I'd suggest adding more explanation of these inputs when you introduce the algorithm.
- Fig. 5:

  - are the times shown clock time or CPU time?
  - while run time is helpful, it can vary drastically depending on implementation, language, system, etc. You may want to consider also including a count of total calls to your objective function.
- Fig. 6, 8:
  - Are your wind turbine markers to scale?
  - This figure is missing axis labels
  - This figure is missing units for the tick labels
- 5.3: the baseline of the percentages given is unclear. Is each percentage given using the last step level as the baseline or the original "incumbant" value?
- Fig. 7: perhaps I missed where this was stated, but are all the AEP values here calculated using the full model for comparison? If not, I think they should be.
- Eq. 23: The equation in your reference is general, but you provide a specific version here. It would be helpful to introduce the general form of your equation from your reference and then fill in the specifics. You may also want to use a more concrete reference here than Investopedia. There are many for this material.
- Line 474: The last sentence here needs more explanation.
- Line 476: I'm not sure what you mean, but if it is the main question then I should. Can you be more specific and/or clarify?
- Fig. 11-13 would probably be more clear if combined and corresponding lines were plotted on the same axes

Technical Corrections

General

- There are many grammar and usage errors throughout. The manuscript should be carefully edited to address these concerns so the material will be more accessible and useful to the community. I have noted a few of these below.

Abstract

1: Introduction

- Line 13: "Subsidy-free . . . " check grammar
- Line 19: Because you give a list of parts here (rather than just one primary thing), "consists of" may be more appropriate.

2: Physics modeling

- Line 132-133: comma after AEP

- Eq. (8): this piece-wise equation contains multiple definitions for some cross-over points. Check the usage of "<" vs "<="

3: Optimization models

- Line 145: check commas to ensure clarity
- Line 200-203: check grammar and usage to ensure clarity
- Line 205: does "this" refer to the linear superposition model or the power-curve free model? In general, try to avoid "this" where there is any possibility of misinterpretation.

4: Neighborhood search heuristic

- Line 243: observation should be singular

5: Computational experiments

- Fig. 4: This figure is a little busy, consider giving the figures a little more space by removing all unnecessary elements and adding some buffer space between sub-figures and figure elements. I really like this figure overall though.
- Line 276: radii
- Line 412: I suggest avoiding starting a paragraph with "As shown in Fig. x" because we don't even know what the subject of the paragraph is yet. The "as shown..." should fit well at the end of the sentence.
- Figures in general:
  - The units given sometimes lead to very large numbers that clutter the figure and impede interpretation. I'd suggest using units that reduce the number of digits required in the tick labels (i.e. GWh instead of MWh, and hours or days instead of seconds)