Review of wes-2024-188

Gaussian process surrogate modeling for efficient controller tuning and fatigue load prediction of the helix wake-mixing method

Authored by: Daan van der Hoek, Tim Dammann, and Jan-Willem van Wingerden

Dear Authors,

I very much enjoyed reading and reviewing your paper. I think it is a very good piece of work, worthy of publication in the *Wind Energy Science* Journal after minor revisions. If the major and minor comments listed below are addressed, I support a publication in WES.

Best regards,

Reviewer

Major comments:

- Your GP surrogate is trained on inflow slices from a single wake. Have you validated that this is sufficient to represent wind farm configurations, where multiple wakes are typically present and interact ? While I understand the useful computational efficiency of your approach, please comment more on the validity to go from single wake to wind farms, and state this as one of the limitation of your study. A validation against wind farm cases (with multiple wakes) would be very interesting to see (I am not asking for it in this paper, but you can state as required future work for improved validation).
- At many occasions, you use symbols without properly defining them. Around an equation or any results, all used symbols should be clearly defined in the text, without ambiguity.
- In addition to text clarifications of symbols, it would be good to add a complete "Nomenclature and abbreviations" section in the appendix, for clarity of all your symbols and acronyms along the paper.
- Fig. 8: It is highly unusual to see negative values for the Strouhal number. In standard aerodynamic, the Strouhal number is defined as a strictly positive quantity. I understand the sign convention that you use in Eq. (1) to separate CW and CCW Helix cases, but this actually applies on your actuated f_{Beta}, not on the Strouhal itself. In Figure 8, I would replace the x-axis with (f_{Beta} – f_{r})*D/U_{inf}. Also in the caption of Figure 8, add clear separation: "comparing the CW (<0) and CCW (>0) helix implementations". Looking at the figure alone, it is not clear where are located CW and CCW for a non-very careful reader (you can even think

of adding CW and CCW above the first plot where it is applicable, if you want your plot to have a clear impact).

- Fig. (9) and Table (4) present results of the GP performances over the whole aggregated dataset (including cases of greedy control, yaw misalignments, and Helix operation), right? It would be very interesting to see how the performances of the surrogate change whether the turbine is being actuated with WFC, or not (so de-aggregate the RMSPE results into cases of greedy, wake steering and Helix-actuated turbine). Then the robustness of the surrogate approach for WFC could better be evaluated. If the results in Fig. (9) and Table (4) include solely greedy-operating turbine in the wake of an upstream actuated turbine, this should be clearly clarified (I am unsure of this information).
- Section 3: After reading the whole paper, it is still unclear (methodologically and results-wise) why do you use GP to predict the optimal control setpoints of the Helix? What is the application of it? A GP should be needed only if the output should be interpolated efficiently in many future different configurations. But in the present case of optimal setpoints selection, you actually yourself refine in the regions that are promising, so in the end your final (fixed?) choice of St = 0.25 is not really coming from the GP, but rather selected directly from a pool of points that you had to compute in AMR-Wind anyway. I do not see the application of the GP in Section 3 for setpoints selection, apart from drawing nice contours. In different ambient wind conditions or wind farm configuration, would the optimal setpoints change, so that you need to apply your GP to find new ones each time? But this GP was trained on this two-turbine setup, so it looks only applicable in the cases that you already ran. So there was no need for a GP, after all. I am being overly critic here to clarify my thought, but please add a general discussion on this point (why is a GP needed at all in Section 3?).
- Are you going to share any data (simulation, trained GP, etc.) publicly with this paper, as it is promoted for Wind Energy Science? This would help to reproduce your findings or directly apply your GP models, and thus would increase the visibility of your work to the community.

Minor comments:

- Eq. (7): your definition of TI is incorrect. You are missing the time-averaging operator (typically denoted as overbar, and explained what it means) over the u'u', v'v', and w'w'. Without this time-averaging, your TI is instantaneous as it is written (which does not really makes sense).
- Line 110: Shouldn't the surface roughness has a unit in AMR-Wind (meters, I expect)?
- Line 112: Which wind speeds exactly? Would be good to specify (in a table for example). Also, what TI level, what vertical shear ? Any veer included ? I see now that (most of) this information is in Table 1, but please add already a pointer to the Table 1 here on line 112, to avoid sounding a bit vague.

- Line 114: what do you mean exactly by "generally" ? In some cases, not?
- Figure 1 caption: Please add temperature (Theta [K]) first in the caption. It's odd to list only three plots when there are four. Also, adapt the y-axis range to be uniform across all 4 plots, for direct comparison and clarity.
- Lines 123-124: please be a bit more precise here where it naturally fits (OpenFAST is only a wrapper code): "the aerodynamic forces of the turbine are modeled using the AeroDyn (v15.03?) module of OpenFAST."
- Line 126 “59 actuator points”: Again here please be a bit more precise shortly. Where are those actuator points located ? Are those on the blades ? Is it 59/3 points per blades (which does not match a integer number)? Clarify with an additional (half-)sentence.
- Line 132: typo OpenFAST (styled capitalized).
- Section 2.2 Wind farm flow control methods: While I agree on keeping this section short and clear to the point as it is, references should be added here and there to point to published literature on both wake steering and the Helix method, to guide the interested reader to go read other solid sources on them if more details are needed.
- Figure 2: Expand this caption with more information on the main elements of the figure. The reader should not have to refer to the main text to understand your figure. Especially, what is the 4Dx10D subdomain here ? Is it the whole simulation domain, or a different mesh refinement level ? Unclear here from the figure + caption. I see (later) that it's explained in the text, but the figure + caption should be self-contained.
- Table 1: Shouldn't the surface roughness has a unit in AMR-Wind (meters, I expect)?
- Lines 142-145: This part should be inverted, to respect the physical order from causes to consequences. From a control perspective, sinusoidal individual pitch references signals are imposed on the turbine, which results in desired varying tilt and yaw moments. The way that you wrote, it sounds like the moments are the origin cause and the pitch signals are the consequence (which is physically wrong, in my understanding).
- Line 145 “the frequency of the pitch signals”: You forget to define that it's f_{Beta} here. Equation 1 does not define what is f_{Beta}, it should be defined in the text.
- Line 145 “Strouhal number”: Again, define here that it's denoted St. Equations (2) is unclear otherwise.
- Line 150 “excitation frequency”: just above you name it actuation frequency, and now here excitation frequency. Please unify for clarity. I think “excitation frequency” is the most correct in both here, so please correct on line 149 before the equation. Actuation should be more the used term for the f_{Beta}.
- Line 150-151: This sentence is unclear, I needed to read it five times. It would be much more straightforward and clear to say: "In case of the counterclockwise helix implementation, the

excitation frequency is added to the rotational frequency, while in the clockwise version it is subtracted from it."

- Figure 3: Here you should more rigorously define the symbols used. What are Theta 1, 2, 3? You haven't defined that it's the individual pitch angle of each blade. Furthermore, here you use a Helix amplitude A_{Beta} of 2° (or 4°?) as example. Later you talk about the cases of A_{Beta} that you considered, but here in this part is missing an equation showing how A_{Beta} is integrated into the pitch signal. All of this might be obvious to you, but not to a reader non-initiated to the Helix method.
- Line 182 "one-second interval": so your wind fields for the surrogate training have a time step of 1s. Is that accurate enough to capture the high-frequency effects on DELs ? (I am doubting). Please comment on this point. After simulating your wind farm in LES at 0.05s timestep (which probably was expensive), wasn't it easily doable to extract the wind slices at higher frequency than 1Hz?
- Line 193-194: you haven't properly defined where/how the amplitude A_{Beta} influences your control actuation (for readers non-initiated to the Helix).
- Figure 4: "Schematic of the load surrogate model": It should rather be "Schematic of the process to train and validate the load surrogate model used in this study". Since the load surrogate model per say is only a subblock in Figure 4.
- Equation 6: what is u(Theta, r) exactly ? You should first define that it's the local time-averaged wind speeds for the streamwise direction. The time averaging is done before the space averaging of Equation 6, I presume.
- Line 207 "This was done to distinguish between the wake of a turbine operated with greedy control and the helix method, as the latter generally causes higher velocity fluctuations.": This explanation is a bit unclear and not really justifying the choice of U_inf. In the wake of a helix-actuated turbine, there is higher velocity fluctuations but also higher mean wind speeds (if the Helix is effective to enhance wake recovery). So if you chose to normalize TI by the local wind speeds instead of U_inf, the values of TI would have been more balanced. I am not saying to change your choice, but clarify the discussion here. If you start an explanation of your choice (which I appreciate), make it complete.
- Table 3: you list the rotor speed as an output, but there isn't any results of rotor speed presented in this paper. I haven't checked for each output channels you list, but I consider that only results that are included in the paper should be listed here (the matter is less of output channels from your simulations, that actual selected channels that you show in the paper, I believe). Furthermore (caught my eye), why do you use the unit s^-1 for rotor speed instead of the most commonly used RPM (in wind energy literature and in OpenFAST)? I would just delete this line (and maybe other) from Table 3 if it's not used in the rest of the paper.

- Line 214 “N_{bins} = 200”: you use the subscript {bins} here but not inside Eq. (8), please unify. Here and everywhere in the paper, check that all symbols are cleanly and unambiguously defined before or where they are used.
- Line 215: “N_{cycles} refers to the number of 1Hz cycles in the 10-minute time series”: Add here “, as per convention,”. This choice of using “1Hz DELs” is a common convention in our domain, but other choices could be valid (some literature use for example 2e6 cycles over 20 years operation, which just scales differently the DELs).
- Line 255: “preprocessing”: shouldn’t it be rather “post-processing”?
- Line 256 “Matern covariance function”: add a reference for it.
- Line 273 “The next step in the modeling framework is to expand the GP model to include pitch amplitude.”: Why did you do it iteratively? Wouldn’t it be more efficient and more accurate for the GP to directly learn on both the Strouhal and the pitch amplitude? Are there good reasons to decouple them as you did?
- Eq. (12) is just a repetition of Eq. (1) and (2). Please just point to the first ones. Also because it’s odd to define equations in your results section 3., it should all be contained in methods section 2.
- Line 315 “averaged over the six turbulence seeds to obtain a steady-state representation of the fatigue loads.”: that goal is unfeasible, per the very definition of fatigue loads (that are time-varying). What you mean is: “averaged over the six turbulence seeds to obtain converged quantities of the fatigue loads, helping towards predicting them from a steady-state representation of the inflow.”
- Line 385 “However, the average trend is similar to that of T1 for most components.”: I think this statement is a bit misleading, since the root causes for T1 and T2 are so different (respectively increased actuations, and more turbulent wake). Furthermore, the statement is not correct for the tower base fore-aft DEL according to your results Fig. 12, where it increases for T1 but decreases for T2.
- Line 393-395 “With AMR-Wind, a pitch…”: this whole last sentence raises an important point without clarifying it enough. Isn’t AMR-Wind using OpenFAST to simulate the turbine aerodynamic and structural behavior? For an upstream turbine, how can the results differ so much? And how can you know which one to trust more? You write “the power loss of T1 is overestimated” but that clearly depends on which result you consider more accurate. Please add clarifications here.