Master Document: Multi-Fidelity BO
Aman Choudhri (ac4972)
December 1, 2024

Contents

1 Definitions and Thoughts

1.1 Useful Definitions

Large eddy simulation (LES). A numerical method to approximately simulate turbulent flows more cheaply than directly solving the Navier-Stokes equations. The key idea behind LES is to essentially ignore the smallest length scales, which add significant computational overhead, via a low-pass filter on the Navier-Stokes equations. Instead, the small length scales are modeled using "subgrid-scale (SGS) models," which are usually so-called "eddy-viscosity models."

Atmospheric boundary layer (ABL). The lowest level of the atmosphere, a region directly influenced by contact with the Earth's surface. Above the ABL, the wind flows parallel to level curves of equal pressure, but within in the flow is far more complex. It appears that there are multiple common "profiles" of ABLs used in modeling: neutral, stable, unstable, convective. Further, it appears that there are two main approaches for modeling the ABL, either a "precursor LES" or a "synthetic" approach [3, 8].

Gravity wave effects.

Thermal stratification.

Reynolds-Averaged Navier-Stokes (RANS) model. A model for resolving "wake turbulence" that is often used in analytical engineering models. It has many constants which must be tuned empirically using real aerodynamic wind farm data [8, pg.2]. Aside from one case known as the "Reynolds Stress model," RANS models generally make a certain isotropy assumption on the turbulence, rendering these models unsuitable for anisotropic flows like the atmospheric boundary layer and turbine wakes [8, pg.3].

Power controller. Seems to be something relating to constant versus variable speed operation. In [3], the authors describe a modeling approach with a constant tip-speed ratio as one with no power controller.

Local thrust coefficient, C_T.

Blockage effects. Mentioned in [2].

1.2 Useful Facts

As of 2014, a "normal separation for turbines on a modern wind farm" is roughly $7-8D$ [8, pg.4].

The most important eddies to model accurately are those that are comparable with or larger than the turbine diameter D, and ones which are most responsible for the transport of mass, momentum, and energy [8, pg.4].

LES are preferable to engineering models because of their ability to “capture the transient evolution of turbulent eddies that are most relevant to wake development and power production” [8, pg.9].

Field data is highly noisy due to the “erratic nature of the atmosphere,” which may compromise the ability of engineering models to deliver reliable predictions in tail regimes like “very large wind farms or inflow angles involving high number of wake interactions.” [8, pg.15].

As the grid resolution of an LES becomes coarser, the fidelity and accuracy of ABS and SGS modeling becomes more and more important [8].

Thermal stratification in the ABL means that “atmospheric gravity waves can be triggered above the ABL” [11].

It seems like we need to be very careful with periodic boundary conditions, because a potential concern is that wakes may be “re-introduced into the domain” [11, pg.303].

“Actuator disk theory tends to overestimate the power output of wind turbines at large yaw angles” [2, pg.878], which actually came from a 2019 Porte-Agel paper. Bempedelis is an advocate for actuator disk theory, nonetheless—the comment was included only as a nod to a reviewer.

“The normalized velocity deficit in the turbine wakes has been observed to follow a self-similar Gaussian profile in several experimental and numerical research studies” [9, pg.2].

1.3 Working Thoughts and Observations

It seems like we only care about modeling the far wake? Maybe we place turbines within $6D$ of each other, but that does seem kind of close...

Essentially the closer the turbines are together, the more accurately we want to model. In section 4.2.2 of [8], the authors describe a study that used the incredibly-expensive but more accurate “actuator line” approach to model turbines since the spacing was roughly $4.3D$, which “would require proper modelling of the near wake.”

ABL modeling and potentially “coupling” with aeroelastic codes could be useful for modeling turbine loading in “off-design conditions like non-neutral ABLs and gusts” [8]. This could maybe be useful if we’re framing this paper as trying to accurately estimate annual energy production across a variety of historical wind conditions.

It seems like there is some kind of “inherent variability of LES,” [3], which reminds me of the phrase “time-averaged.” Is there some kind of literature on the length of time one runs an LES for to determine a power production for a given wind speed? It seems like [1] discusses this in detail.

It seems like the idea of ABL modeling is because “wind turbine wake recovery and thus power production are greatly influenced by background atmospheric turbulence” [11, pg.302].

FRAMING: Loosely motivated by the results observed in [2], where a wake model gradient-based optimization routine produced better results than an LES simulation in 30% of runs, I’m curious about the idea of framing this paper in terms of LES checking the results of wake models. Especially since the actual efficiency predictions from wake models was so drastically different across FLORIS and LES, again in [2]

TODO: Let's perform a detailed review of papers comparing the performance of wake models and LES simulations.

1.4 Question Log

1.4.1 Research Directions

1. Can we model the effects of downstream dynamic loading on turbines and their longevities? What have people done in the past? Can we formulate this as a multi-objective optimization problem? Or by reducing wake effects in general do we solve this inherently?
2. Can we take advantage of our knowledge about the kinds of eddies that are most important to model correctly to specify more carefully the relationship between multiple fidelities of wake models? Ie if we know that an analytical approximation underestimates wake performance in a certain regime, we can get a better sense for how the resulting observation might be biased?
3. Can we encourage the model to tune different parameters of the LES or use different approximations based on the specific input case? Or is that way too much effort?
4. On this note of tuning different parameters, can we take advantage of what we know about the physical model to understand the BIAS or NOISE of certain observation fidelities in different input regions? I.e. we might expect higher noise or a certain bias direction in the LES model based on a function $g(X)$ defined on our configuration space, which returns the average separation between turbines?
5. On this further note, is it worth also incorporating a RANS fidelity? I guess to understand this, my main questions would be: how well does RANS perform compared to LES? Like what are the real main disadvantages? And what are the real computational advantages?
6. Can we take advantage of interesting *mixed* approximations like what is described in section 4.3 of [8], which uses LES to improve simpler models? Some keywords to look into here include a "modified Frandsen's model" and an "LES generated transfer function."
7. Is it worth incorporating different rotor dimensions? If I go the TURBO route and try to make the case that BO is great in high-dimensional settings, it could make sense. It could also add additional complexity and make it more difficult to do anything interesting relating to the GP kernel.
8. Can we use a multi-output GP that gives us the wind speed at each turbine location as our model? And then somehow average over the power productions at each location in our acquisition function? This way we can maybe more carefully specify the bias or variance of a given observation fidelity $f^{(m)}$?
9. Does it make sense to think about running multiple LESs in parallel as well? Can we run one LES and one RANS in parallel? Or should we cap it at one and use all available cores for parallelization of the actual LES algorithm?
10. Can we adaptively select the temporal averaging window for a given simulation to result in various levels of observation noise?

1.4.2 General/Empirical Questions

1. What is this concept of “spanwise” versus “streamwise” spacing? Does it refer to different behavior in terms of the spacing of turbines in directions parallel to versus perpendicular to the wind inflow direction? How do existing BO studies handle this distinction constraint-wise, if at all?
2. Should we use pseudo-spectral or finite-volume approaches? What are the advantages and disadvantages of each?
3. What are the key ideas behind different wake models and their assumptions? Why do people use the GCH model [9]? Does it just empirically work very well?
4. What is the actual variability of the simulations in a 50-turbine setting? Can we get away with 10-minutes of wind farm time? Can we bound the variance of the estimations in those contexts? Is there a burn-in period?

2 Literature Review

2.1 Bayesian Optimization Methodology

Javier Gonzalez et al. “‘Batch Bayesian Optimization via Local Penalization’”. *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*. 2016.

- **Overview:** Proposes a new batch BO method based on a Lipschitz assumption on the objective function. Uses the GP to infer the Lipschitz constant. Greedily selects points within a batch by penalizing the acquisition function away from previous points in the batch. They base their penalization on the estimated Lipschitz constant—specifically, the width of the downweighted region around a previously-selected point $x_{t,i}$ is based on how close $x_{t,i}$ is believed to be to the optimum, which itself is derived from an estimate of the Lipschitz constant.
- **Strong points:**
  – Gradient of the penalized acquisition function available in closed form.
- **Weak points:**
  – Assumes that the kernel function K is twice-differentiable, in order to estimate the Lipschitz constant.
  – Assumes that the function is Lipschitz homoskedastic, i.e. that the Lipschitz constant of the function is the same everywhere. The authors note that the method can be extended to cases where a local Lipschitz constant is estimated, but they don’t expand on it.
- **Empirical results:**
- **Takeaways:**

2.2 Wind Farm Optimization

Nikolaos Bempedelis et al. "'Data-Driven Optimisation of Wind Farm Layout and Wake Steering with Large-Eddy Simulations'". *Wind Energy Science*. 2024.

- **Overview:** Applies Bayesian optimization with LES and analytic wake models to maximize power production in both micro-siting and wake steering tasks. Doesn't do anything truly multi-fidelity.
- **Methodological details:**
  - Smagorinsky SGS model and AD-NR method used, simulated using Winc3D package [5].
  - Precursor simulations of "pressure-gradient-driven fully developed neutral" ABLs.
  - Enforces rotor spacing of D. Sets up a $18D \times 18D$ site, with $N = 16$ turbines. Picks a $D = 100$m rotor diameter and $h = 100$m hub height. Assumes a constant local thrust coefficient $C_T' = 4/3$.
  - *BO parameters*: No observation noise. Lower confidence bound acquisition function. Kernel hyperparameter search and acquisition function maximization carried out with L-BFGS algorithm. Batching with "local penalization," as described in [6].
  - Evenly-weighted 6-directional wind rose.
  - Flow data are averaged over a 2.5h period of farm operation (pg.873).
  - They compare their framework, LES-BO, against a gradient-based optimization strategy using the "Gauss-curl hybrid" wake model from the FLORIS package. The gradient-based optimization strategy relies on what appears to be a convergence tolerance threshold, which means it's unclear how many evaluations the FLORIS strategy had compared to the LES-BO model. A good question to ask the authors.
  - Simulation costs ranged between 300-900 CPU hours, performed in parallel on 128 or 256 cores.
- **Empirical results:**
  - Finds a rational quadratic kernel to outperform both a Matern and squared exponential, where by performance they mean extrapolation performance on heldout test points.
  - They find that roughly 30% of the layout designs found by the analytical wake model outperform their solution using BO and the LES.
  - They also do find that LES outperforms FLORIS in the wake-steering task, however.
  - FLORIS underpredicts efficiency compared to the true efficiencies calculated by LES.
- **Takeaways:** The performance from the FLORIS wake model simulation is maybe not the best sign for the importance of including LES observations. To truly understand the cause of the performance difference, however, I want to go into their Appendix B and understand how many samples and how many evaluations were used for each. Because if the FLORIS approach has far more evaluations available, this is maybe not a fair comparison. Could be that optimizing gradients over the analytical wake model is likely to result in many local optima.

Tinkle Chugh and Endi Ymeraj. "'Wind Farm Layout Optimisation Using Set Based Multi-Objective Bayesian Optimisation'". *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. 2022.

- **Overview:** Applies multi-objective BO with objectives of wind farm power output and turbine cost, using expected hypervolume improvement as the acquisition function. Formulates the problem as a *set-valued* search space, defining a custom kernel over collections of turbine location coordinates and enabling search over variable numbers of turbines.
- **Strong points:**
  - Allows for *variable numbers of turbines*, with a novel set kernel function
  - Considers *variability in wind speed and direction*. Models power output as an expectation over $p(v, \theta)$, a joint distribution over wind speed and direction taken from historical data, estimated using Kernel Density Estimation.
- **Weak points:**
  - Only considers the Jensen model, with *no large-eddy simulations*.
  - Discretizes the space of possible turbine locations in the acquisition function to a 20x20 grid.
- **Interesting details:**
  - Maximizes EHVI using a genetic algorithm.
  - Sets minimum distance between turbines as 3 times rotor diameter.

Daan van der Hoek et al. "*Predicting the Benefit of Wake Steering on the Annual Energy Production of a Wind Farm Using Large Eddy Simulations and Gaussian Process Regression*". 2020.

- **Overview:**
- **Strong points:**
  - Incorporates a *time-varying* wind direction, defining power output as the weighted average over the mean direction μ, as well as $\mu \pm 3°$ and $\mu \pm 6°$
- **Weak points:**
  -
- **Interesting details:**

2.3 LES Surveys

D. Mehta et al. "'Large Eddy Simulation of Wind Farm Aerodynamics: A Review'". *Journal of Wind Engineering and Industrial Aerodynamics*. 2014.

- **Overview:** Summary of various LES implementations to model wind farm aerodynamics. Touches on a discussion for how to optimally use LES and challenges of such models.
- **Details:**
  - So-called engineering models are simulations using "basic principles of physics and empirically established approximations." They cannot account for phenomena like "wake meandering...a turbine's response to partial wake interaction..." etc. Generally, the models used in practice focus on modeling the *far-wake*, and many resolve turbulence using Reynolds-averaged Navier-Stokes models, which have constants that need to be empirically specified using real aerodynamic data.
  - The two key factors of a wake from a turbine are the "velocity deficit," the reduction in wind velocity due to the energy extracted by the turbine, and the "added turbulence

intensity," the increase in turbulence of the flow within a wake. The velocity deficit reduces the power that downstream turbines can generate. The added turbulence intensity increases the "dynamic loading" of downstream turbines, reducing their longevity.

- Generally the wake of a turbine is considered as two regions, near and far. The near wake is immediately behind the turbine, and the key factor is the velocity deficit, which "attains its maximum value between $1D$ and $2D$," where D is the turbine rotor diameter. In this region, the key influencing factors are the design of a turbine and its loading (what?). The near wake generally ends between $2D$ and $4D$.
- The velocity deficit is generally negligible beyond $10D$ but increased turbulence intensity is sensible up to at least $15D$.
- Wind turbine and wake flow modeling is a regime well-suited for incompressible Navier-Stokes. "Widely accepted that the incompressible form of the NS equations can model the flow around a wind turbine and in its wake." (page 2)
- Wind turbine wakes are high-Reynolds number flows, which make direct numerical simulation using Navier-Stokes infeasibly expensive. This is because higher Reynolds number flows means shorter and shorter length scales of eddies, meaning the overall range of possible eddy scales grows larger. (page 3)
- The main paradigm for studying turbulence involve some kind of statistical averaging, removing certain length and time scales. But the problem is underdetermined, known as the "closure problem," leading to different mathematical models for the flow. This is the origin of RANS and LES.
- The Reynolds Stress model, which does not make the isotropy assumption of most RANS models, "provides the most reliable results in both the near and the far wake," among far wake models. It does underpredict velocity deficit in the near wake, however (page 4).
- Engineering models perform poorly for a single "inflow angle" but do better when averaged across multiple angles. The poor single direction performance "stems from the combination of innacurate turbulence and ABL modelling, tuning with limited field data, and the inability to resolve the flow in case of multiple turbine-wake or wake-wake interactions."
- Argues that industry needs "detailed knowledge on the performance of wind turbines post-deployment" in "off-design" regimes, like with gusts or atmospheric stratification.
- A common class of "eddy-viscosity" SGS models, those derived from "Smagorinsky's model," make a certain assumption on the alignment between the strain rate and subgrid tensors, which is unsupported by numerical simulation data. Other models like the "Scale Similarity Model" and derivatives like the "Mixed SSM" do not make this assumption.
- Wind turbines are usually modeled using *actuator methods* instead of direct methods, which are computationally expensive.
- "Energy-conserving schemes" are flagged as promising, being free from "numerical dissipation" and the requirement to use "periodic boundaires," but are noted as requiring further study. I'm curious whether these have been looked into further as of 2024.
- Another open-ish question from this paper is the best way to handle boundary conditions with the ABL. Not sure what this means, but a keyword is "Monin-Obukov's approach," which is flagged as being "unsuitable for LES." Other approaches are briefly alluded to, but the authors note that more experiments are necessary.
- Empirical results:

        * The AD-NR actuator disk method is "suitable and faster" compared to the more expensive AD-R method (which incorporates "tangential forces") if "one aims to analyze only the power produced and if turbines are separated by at least 5D to 7D" (pg.4)

S.-P. Breton et al. "'A Survey of Modelling Methods for High-Fidelity Wind Farm Simulations Using Large Eddy Simulation'". *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2017.

- **Overview:** Survey of common schemes for modeling rotor, atmospheric conditions, and terrain effects within LES implementations as of 2017. Not as focused on the CFD methods. Also summarizes experimental research data available for validating LES implementations.
- **Details:**
    - Argues that the most practical concern for wind farm simulations is to understand the ways that "wake effects alter turbine loading and power extraction and how such effects are influenced by atmosphereic conditions and topography."
    - Alternatives to LES models include "engineering models" (analytic wake models?), computations solving the Reynolds-averaged Navier-Stokes (RANS) equations, and direct numerical simulations. Direct numerical simulations are too expensive to use, they argue. And apparently, RANS-based models are "known to depend on the choice of turbulence closure models" and they can only provide limited information about the "inherently unsteady processes underlying wake phenomena" because of their time-averaged formulation.
    - There are multiple ways to computationally solve the filtered Navier-Stokes equations within LESs, including finite difference, finite volume, or pseudo-spectral.
    - Supports the idea of including some component of aeroelastic coupling. "Considering aeroelastic effects is essential when simulating multi-megawatt turbines; the assumption of small blade deformations...loses accuracy with increasing blade size." (pg.15). Keyword: FAST for the coupled aeroelastic model.
    - Advocates for AD+R approaches to modeling turbine rotor as a good balance of cost-effectiveness for modeling far wakes and wake interactions. Probably the regime we're in for power production modeling.
    - Notes the importance of the use of a "power controller to actively determine the rotational speed of the rotors."
    - There is no definitive best approach between synthetic and precursor LES modeling for the ABL.

2.4 LES Implementations

Georgios Deskos, Sylvain Laizet, and Rafael Palacios. "'WInc3D: A Novel Framework for Turbulence-Resolving Simulations of Wind Farm Wake Interactions'". *Wind Energy*. 2020.

- **Overview:**
- **Details:**

  - Finite-difference discretization.
  - Uses the Smagorinsky SGS model, with some Mason-Thomson correction (not sure if this is standard or not).
- **Claims:**

S. Stipa et al. "'TOSCA – an Open-Source, Finite-Volume, Large-Eddy Simulation (LES) Environment for Wind Farm Flows'". *Wind Energy Science*. 2024.

- **Overview:** Open-source, finite-volume LES aimed at large-scale studies of wind farm-induced gravity waves and the interaction between cluster wakes and the atmosphere. Specifically designed for LES of large finite wind farms.
- **Details:**
  - Supports actuator line, ALM, actuator disk, ADM, and uniform actuator disk, UAD methods. Presumably UAD is AD-NR, and ADM is AD-R.
  - Finite-volume framework.
  - Supports a "concurrent-precursor method," which tye argue may not be available in other finite-volume solvers (even though it is "extensively used in pseudo-spectral methods.").
  - Supports a "sharp-interface immersed boundary method (IBM)" that they argue allows for the simulation of moving objects and complex terrain features.
  - Enforces a "desired hub-height wind speed" while avoiding "inertial oscillations produced by the Coriolis force above the boundary layer." No idea what this means.
- **Claims:**
  - Notes that industry primarily uses analytical, "reduced-order wake models" to estimate annual energy production. Cites [10] to argue that these models struggle with reproducing wind farm blockage and farm-farm wake interactions.
  - Notes that "only a few" existing LES implementations can tackle "gravity wave effects."
  - They argue that finite-volume approaches, by virtue of allowing for "grid stretching," enable the resolution of larger domains with the same number of degrees of freedom and also providing more "geometric flexibility."

References

[1] S. J. Andersen et al. “Quantifying Variability of Large Eddy Simulations of Very Large Wind Farms”. In: *Journal of Physics: Conference Series* 625.1 (June 2015), p. 012027. ISSN: 1742-6596. DOI: 10.1088/1742-6596/625/1/012027. (Visited on 11/24/2024).

[2] Nikolaos Bempedelis et al. “Data-Driven Optimisation of Wind Farm Layout and Wake Steering with Large-Eddy Simulations”. In: *Wind Energy Science* 9.4 (Apr. 2024), pp. 869–882. ISSN: 2366-7451. DOI: 10.5194/wes-9-869-2024. (Visited on 11/15/2024).

[3] S.-P. Breton et al. “A Survey of Modelling Methods for High-Fidelity Wind Farm Simulations Using Large Eddy Simulation”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375.2091 (Mar. 2017), p. 20160097. DOI: 10.1098/rsta.2016.0097. (Visited on 11/24/2024).

[4] Tinkle Chugh and Endi Ymeraj. “Wind Farm Layout Optimisation Using Set Based Multi-Objective Bayesian Optimisation”. In: *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. Boston Massachusetts: ACM, July 2022, pp. 695–698. ISBN: 978-1-4503-9268-6. DOI: 10.1145/3520304.3528951. (Visited on 11/12/2024).

[5] Georgios Deskos, Sylvain Laizet, and Rafael Palacios. “WInc3D: A Novel Framework for Turbulence-Resolving Simulations of Wind Farm Wake Interactions”. In: *Wind Energy* 23.3 (2020), pp. 779–794. ISSN: 1099-1824. DOI: 10.1002/we.2458. (Visited on 11/19/2024).

[6] Javier Gonzalez et al. “Batch Bayesian Optimization via Local Penalization”. In: *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*. PMLR, May 2016, pp. 648–657. (Visited on 11/24/2024).

[7] Daan van der Hoek et al. *Predicting the Benefit of Wake Steering on the Annual Energy Production of a Wind Farm Using Large Eddy Simulations and Gaussian Process Regression*. Mar. 2020. arXiv: 2003.12153. (Visited on 11/16/2024).

[8] D. Mehta et al. “Large Eddy Simulation of Wind Farm Aerodynamics: A Review”. In: *Journal of Wind Engineering and Industrial Aerodynamics* 133 (Oct. 2014), pp. 1–17. ISSN: 0167-6105. DOI: 10.1016/j.jweia.2014.07.002. (Visited on 11/24/2024).

[9] Amin Niayifar and Fernando Porté-Agel. “Analytical Modeling of Wind Farms: A New Approach for Power Prediction”. In: *Energies* 9.9 (Sept. 2016), p. 741. ISSN: 1996-1073. DOI: 10.3390/en9090741. (Visited on 11/12/2024).

[10] N. G. Nygaard et al. “Large-Scale Benchmarking of Wake Models for Offshore Wind Farms”. In: *Journal of Physics: Conference Series* 2265.2 (May 2022), p. 022008. ISSN: 1742-6596. DOI: 10.1088/1742-6596/2265/2/022008. (Visited on 11/24/2024).

[11] S. Stipa et al. “TOSCA – an Open-Source, Finite-Volume, Large-Eddy Simulation (LES) Environment for Wind Farm Flows”. In: *Wind Energy Science* 9.2 (2024), pp. 297–320. DOI: 10.5194/wes-9-297-2024.