Review of

Stochastic Gradient Descent for Wind Farm Optimization

Julian Quick et al.

Reviewer:

Ahmad Vasel-Be-Hagh

Fluid Mechanics Research Laboratory, Mechanical Engineering Department, Tennessee Technological University

The article describes using Stochastic gradient descent (SGD) (it is misspelled as "decent" in line 27) to optimize the layout of a wind farm with the objective of maximizing annual energy production (AEP). SGD is a pretty old algorithm; however, the authors used a modified SGC called Adam SGD, developed in 2014.

My primary concern with this article is the novelty. This paper is a single-objective layout optimization that compares the AEP production and computing time for the said algorithm and a second one serving as the baseline. The second algorithm that the article uses as its baseline is Scipy Sequential Least Squares Programming. The two algorithms lead to layouts that generate almost equal AEP (Adam SGD produces ~ 0.2% more, which is not significant). The computational time for the Adam SGD algorithm is nearly 0.5 hours. In comparison, Scipy Sequential Least Squares Programming takes almost 10 hours to identify an optimal layout. This data is for the 100-turbine case, which yields the maximum difference among all presented instances. As the authors also acknowledge, the Scipy Sequential Least Squares Programming is a 30-year-old Fortran code, which one might be able to modify to make it more efficient. Also, other modern algorithms can generate equally efficient layouts as quickly as Adam SGD. Hence, showing that Adam SGD finds layouts that are approximately as efficient as those found by another specific algorithm but does so more quickly does not offer enough novelty, in my opinion. In addition, the industry and the public, including investors, are not concerned with spending several more hours or days finding more efficient algorithms since they invest hundreds of millions of dollars in a wind farm. Speeding up the optimizations is crucial for real-time control of wind farms, not layout optimization, since that is only a one-time job. With that said, again, I must mention that this article does not demonstrate that Adam SGD is the fastest optimizer; it shows that Adam SGD is faster than one other algorithm, leaving the reader to believe there are other algorithms that are faster than Adam SGD.

Below are my other comments on this article:

My other concern is about the spacing constraint: "*We require each turbine to be spaced at minimum two rotor diameters apart (ND = 2)*" [Lines 163-164] But then, the article reads: "*All power plants considered in this study have square boundaries so that, in a grid, the turbines would be spaced five rotor diameters apart (Δ = 5).*"
These statements confuse me. Is the spacing 2 or 5 diameters? This matters since ND = 2 is exceptionally tight, and I am not aware of any real wind farm with turbines that are so close. Lillgrund, of which the authors used its wind data for this work, is considered a packed wind farm, in which I believe the two closest turbines in one specific direction are yet apart by more

than 4D. Also, what do authors mean by "in a grid?" Did they define a background grid to discretize the search area or the search area was continuous?

Line 20: "It can be costly to compute the power associated with each speed and direction combination, as the number of these combinations can be large." -
I don't believe these computations are costly, and a modern processor can quickly handle such calculations. Also, the wind speed and direction are measured using sensors with limited resolution and other uncertainties, and that's why they are reported in a discretized way using bins.

Lines 23: "to avoid discretizing the input distributions into evenly spaced intervals. In this study, we present an approach for wind farm optimization that estimates the gradient of the AEP using Monte Carlo simulation. This does not require that the input be discretized at all, and allows for the consideration of an arbitrarily large number of atmospheric conditions".
Input data are measured in a discretized way. The guide to meteorological instruments and methods of observation recommends measuring wind direction with an accuracy of 5 degrees. Hence, why would feeding the data that is measured discretized in the first place in a "continuous" format into a model help the accuracy?

WMO (World Meteorol. Organ.). 2008. Guide to Meteorological Instruments and Methods of Observations. WMO No 8. Geneva: WMO. 7th ed.

Figure 1: I do not think this figure reflects a realistic representation of Lillgrund's wind conditions. I know that the westerly and southwesterly winds are stronger at this site; however, the data presented in this figure is too extreme. I think this is (at least partly) caused by what I explained in the two previous points. This requires the authors to explain precisely how they converted the data measured discretized to a continuous format. It has to be done by some sort of interpolation. How did they conduct this interpolation? Does the data presented in this figure even sum up to 1 (100%), since the largest frequency shown is about 0.005?

Line 40: "This term introduces sudden steep gradients in the optimization space, necessitating the use of smaller momentum parameters than are typically employed in the Adam SGD algorithm."
The article presumes the reader knows the used algorithm; hence, it employs undefined terms. For instance, in the statement above, what are the momentum parameters, and why/how would smaller ones help? This is not defined by this line of the article.

The introduction does not cover the literature on "wind farm layout optimization (hereafter WFLO)." I also find the introduction scattered and distracted with information that does not belong to the introduction, in particular, or this paper, in general. For example, I think the quick review given on lines 59-61 on applications of SGD in areas other than WFLO is distracting. In these applications, SGD tunes a neural network's hyperparameters. Instead, the introduction needs to remain focused on the literature of WFLO, identify gaps in WFLO, and clarify how this new optimization using SGD addresses the identified gaps.