Wind Farm Layout Optimisation using Set Based Multi-objective Bayesian Optimisation

Tinkle Chugh
t.chugh@exeter.ac.uk
Department of Computer Science
University of Exeter
Exeter, United Kingdom

Endi Ymeraj
ey238@exeter.ac.uk
Department of Computer Science
University of Exeter
Exeter, United Kingdom

ABSTRACT

Wind energy is one of the cleanest renewable electricity sources and can help in addressing the challenge of climate change. One of the drawbacks of wind-generated energy is the large space necessary to install a wind farm; this arises from the fact that placing wind turbines in a limited area would hinder their productivity and therefore not be economically convenient. This naturally leads to an optimisation problem, which has three specific challenges: (1) multiple conflicting objectives (2) computationally expensive simulation models and (3) optimisation over design sets instead of design vectors. The first and second challenges can be addressed by using multi-objective Bayesian optimisation (BO). However, the traditional BO cannot be applied as the optimisation function in the problem relies on design sets instead of design vectors. This paper extends the applicability of multi-objective BO to set based optimisation for solving the wind farm layout problem. We use a set-based kernel in Gaussian process to quantify the correlation between wind farms (with a different number of turbines). The results on the given data set of wind energy and direction clearly show the potential of using set-based multi-objective BO.

CCS CONCEPTS

• **Theory of computation** → **Gaussian processes**; **Mathematical optimization**; • **Mathematics of computing** → **Bayesian computation**.

KEYWORDS

Surrogate modelling, Gaussian process, Renewable Energy, Uncertainty quantification, Gaussian Process Over sets, Pareto optimality

1 INTRODUCTION

Climate change has become a primary concern that needs to be promptly addressed. With the desire to implement its commitments under the Paris Agreement [1], the European Union has proposed a new set of targets for 2030. One of the renewable energies that has increased in popularity is wind energy. This is due to the fact that wind energy is amongst the cleanest source of electricity given the very low greenhouse gas emission and low water consumption [7]. One of the drawbacks of wind-generated energy is the large space necessary to install a wind farm; this arises from the fact that placing wind turbines in a limited area would hinder their individual productivity and therefore not be economically convenient. Because of this, the optimisation of the layout plays an important role in energy yield. The complexity of this optimisation problem is given by the computational power required to reproduce accurate simulations of the wind across a wind farm. Furthermore, the objective function evaluations rely on the design sets instead of design (or decision) vectors. For instance, power is an output of the whole wind farm and not of one turbine. Therefore, to design an optimal wind farm, we need to consider the correlations between wind farms, which can have different number of turbines. In this paper, we use a set-based kernel in Gaussian process model [8] and embed it in the multi-objective BO. To the best of our knowledge, set-based multi-objective BO has never been used to solve a wind farm layout optimisation problem. To be summarised, the contributions of the paper are:

(1) Handling computationally expensive multiple conflicting objectives by using multi-objective BO.
(2) Handling correlations between different wind farms (as different design sets) by using a set based kernel.

The rest of the article is structured as follows. In Section 2, we provide an overview of the wake model and define the optimisation problem. In Section 3, we explain the multi-objective BO using Gaussian process with set based kernel. In Section 4, we provide the results and finally we conclude and provide future research directions in Section 5.

2 WIND FARM LAYOUT OPTIMISATION

2.1 Wake model

The wake effect considers the wind speeds facing different turbines in a wind farm. The wind speeds for different turbines can be different and depend on the coordinates of turbines and the incoming wind speed and direction. A simplified well-known wake models is Jensen model [14], which can be used to estimate the wind speed for each turbine. This model allows for fast computations of the wake effect of turbines since it assumes that the wake behind the rotor expands linearly and that it is only affected by the distance from the turbine. An example to estimate the wind speed v_1 at a distance D from a turbine is shown in Figure 1. In the figure, v_0, r_0 and r_1 are downstream wind speed, turbine rotor radius and radius of cone, respectively. The radius of the cone is: $r_1 = r_0 + \alpha D$, where α is the decay constant and determines the expansion of wake with distance and can be estimated with an analytical formula: $\alpha = \frac{0.5}{\log(\frac{z}{z_0})}$, where z is height of the turbine and z_0 is the surface roughness of the wind farm. In many cases, the α is kept constant. After applying the conservation of momentum, the reduced wind speed v_1 is given by: $v_1 = v_0 \left[1 - \frac{1-\sqrt{1-C_T}}{(1+\alpha\frac{D}{r_1})^2}\right]$, where C_T is the thrust coefficient.

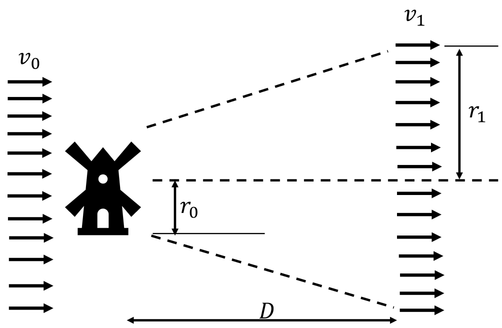

Figure 1: Visual representation of the wake effect from the Jensen's model.

2.2 Objective Functions

One of the main objectives in designing a wind farm is to maximise the power output. In addition, the cost of installing turbines needs to be minimised. The power output depends on the distribution $p(v, \theta)$ of wind speed (v) and direction (θ) and the wind speed at different turbines. The distribution $p(v, \theta)$ can be modelled with the historic data of wind speed and direction. In this way, the uncertainty in wind speed and direction can be handled and quantified in estimating the power output. In this work, we used an open source data set provide by Engie [6] from 2013-2016.

The input to the wake model is the incoming wind speed, direction and location of wind turbines in the wind farm and the output is the wind speed at different turbines. This wind speed is then used to estimate the power curve using the following equation [9]:

$$P_{curve}(v, \mathbf{x}) = \begin{cases} a \cdot \frac{1+m\cdot\exp(-v/\tau)}{1+n\cdot\exp(-v/\tau)} & if \ \ v \geq v_{cut\_in} \\ 0 & elsewhere \end{cases} \quad (1)$$

where (a, m, n, τ) are the parameters (values are taken from [9]) and v_{cut_in} is the wind speed at which wind turbine becomes productive. The $\mathbf{x}$ represents the coordinate or location of turbines and is an input to the wake model. Once the joint distribution of wind speed and direction and wind speed at different turbines is known, the power output of wind farm is defined as [10, 11]:

$$\begin{aligned} P_{tot}(X) &= \sum_{x\in X}\sum_{\theta=0}^{359}\sum_{v=0}^{v_{max}} P_{curve}(v, \mathbf{x})p(v, \theta) \\ P_{tot}(X) &= \text{Expected power of wind farm X} \\ \theta &= \text{Wind direction} \\ v &= \text{Wind speed} \\ |X| &= \text{Cardinality of set X} \\ &= \text{(or number of turbines in the wind farm)} \\ p(v, \theta) &= \text{Joint distribution of wind speed and direction} \end{aligned} \quad (2)$$

As can be seen in the equation above, the expected power of the wind farm depends on the number of turbines. Two wind farms with different number of turbines may give two different output energy. Therefore, to model the objective function, we need to find the correlation between wind farms, which can have different number of turbines. The second objective is the cost, which can be computed as non-linear function of number of wind turbines [12]:

$$\text{cost}(X) = |X| \cdot \left(\frac{2}{3} + \frac{1}{3} \cdot \exp(-0.00174 \cdot |X|^2)\right) \quad (3)$$

The cost as the objective function depends only on the number of turbines and does not use computationally expensive wake simulation models.

Algorithm 1 Multi-objective Bayesian Optimisation

Input: Data Set $\mathcal{D} = \{(X, Y)|X \in \mathbb{R}^{N\times d} \ \ Y \in \mathbb{R}^{N\times w}\}$ and $\mathcal{F}(\mathbf{x}) = (f_1(\mathbf{x}), ..., f_w(\mathbf{x}))$
Output: Evaluated solutions

while Termination criterion is not met **do**
  Fit $\mathcal{GP}_i \ \ \forall \ \ i \in [1, 2, ..., w]$
  Find $\mathbf{x}'$ after maximising the acquisition function α_{EHVI}
  Expensively compute $\mathbf{y}' = \mathcal{F}(\mathbf{x}')$
  Append $\mathbf{x}', \ \mathbf{y}'$ to $\mathcal{D}$

3 MULTI-OBJECTIVE BO OVER SETS

In Multi-Objective BO, we start with a data set: $\mathcal{D} = \{(X, Y)|X \in \mathbb{R}^{N\times d}, Y \in \mathbb{R}^{N\times w}\}$, where d is the number of variables, w is the number of objectives and N is the number of instances in the data. With this data, we want to fit a Gaussian Process ($\mathcal{GP}$) model that helps predict objective functions without direct evaluation. In multi-objective BO, there are two approaches to build surrogate models. The first consists of building a $\mathcal{GP}$ on a scalarized version of objective functions [2]. The second consists in building a Gaussian Process for each objective[4]. This latter option is what will be used in this paper. Once we have fitted the surrogate models we aim to use them to search for a new decision vector ($\mathbf{x}' \in \mathbb{R}^{1\times d}$) that maximises an acquisition function. In this work, we will be using the expected hypervolume improvement (EHVI) [5]. We compute an expensive evaluation ($\mathbf{y}' = \mathcal{F}(\mathbf{x}')$), where $\mathcal{F} = (f_1(\mathbf{x}), ..., f_w(\mathbf{x}))$ and $\mathbf{y}' \in \mathbb{R}^{1\times w}$ and add $(\mathbf{x}', \mathbf{y}')$ to the existing data and repeat the process. The optimisation finishes when a predefined termination criterion is reached. An outline of the algorithm is shown in Algorithm 1.

3.1 Gaussian Processes Over Sets

A $\mathcal{GP}$ can be described as a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix K [13]: $f \sim \mathcal{N}(\boldsymbol{\mu}, K)$. The mean $\boldsymbol{\mu}$ and covariance matrix K are both dependent on a kernel function which quantifies the correlation between instances in the data. For simplicity in calculations, we assume the zero mean. In this work, we use a RBF (also known as squared exponential or Gaussian kernel):

$$k(\mathbf{x}_1, \mathbf{x}_2, \Theta) = \sigma^2 \exp\left(-\frac{||\mathbf{x_1} - \mathbf{x_2}||^2}{2l^2}\right) + \sigma_n^2 \delta_{\mathbf{x}_1, \mathbf{x}_2},$$

where $\Theta = (l, \sigma, \sigma_n)$ is the vector of hyperparameters to be estimated when building the model, $||\mathbf{x_1} - \mathbf{x_2}||^2$ is the squared Euclidean distance between two decision vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ and $\delta_{\mathbf{x}_1,\mathbf{x}_2}$ is the Kronecker delta function. The hyperparameters can be estimated by maximising the following likelihood function:

$$p(\mathbf{y}|\mathcal{D}, \Theta) = \frac{1}{\sqrt{|2\pi K|}} \exp(\mathbf{y}^T K^{-1} \mathbf{y}),$$

Once the optimal parameters are found, we can predict the value of the objective function of a new instance $\mathbf{x}'$ using the following posterior predictive distribution:

$$p(\mathbf{y}'|\mathbf{x}', \mathcal{D}, \Theta) = \mathcal{N}(\mathbf{k}(\mathbf{x}', X)K^{-1}\mathbf{y}, k(\mathbf{x}', \mathbf{x}') - \mathbf{k}(\mathbf{x}', X)K^{-1}\mathbf{k}(\mathbf{x}', X)),$$

where $\mathbf{k}(\mathbf{x}', X)$ is the covariance vector between $\mathbf{x}'$ and the training data X. To deal with the problem of finding the correlation between sets, we utilise Gaussian Process Over Sets [8] which is based on the idea of using a set of decision vectors. In wind farm layout optimisation, the x-coordinate and y-coordinate are the decision variables and therefore, $d = 2$ and the number of turbines n vary for different wind farms. The $\mathcal{GP}$ Over Sets [3] requires adjustments to the traditional Gaussian Process algorithm, namely the way that the correlation between sets is calculated. We compute correlation between sets as follows:

$$k_{set}(X_1, X_2) = \frac{1}{|X_1||X_2|} \sum_{\mathbf{x}_i \in X_1} \sum_{\mathbf{x}_j \in X_2} k(\mathbf{x}_i, \mathbf{x}_j)$$

Two important advantages of using the kernel over sets mentioned above is the resulting covariance matrix is positive-definite and the order of the decision vectors in the set do not effect the correlations between sets. The data set for building the model with N sets is:

$$\mathcal{D} = \{\mathcal{X}, Y \,|\mathcal{X} = [X_1, X_2, ...X_N] \text{ with } X_i \in \mathbb{R}^{|X_i| \times d}, Y \in \mathbb{R}^{N \times w}\}$$

The posterior predictive distribution becomes:

$$p(\mathbf{y}'|X', \mathcal{D}, \Theta) = \mathcal{N}(\mathbf{k}_{set}(X', \mathcal{X})K^{-1}\mathbf{y}, \quad k_{set}(X', X') - \mathbf{k}_{set}(X', \mathcal{X})K^{-1}\mathbf{k}_{set}(X', \mathcal{X})),$$

where $K_{i,j} = k_{set}(X_i, X_j)$ and $\mathbf{k}_{set}(X', \mathcal{X})$ is the correlation vector between the new set X' and the training data $\mathcal{X}$. This formulation is compatible with the objective function formulation that is used in this paper since it allows to compute the correlation of instances with different lengths, where the length corresponds to the number of decision vectors in a set.

To have the same number of decision variables in maximising the acquisition function, we impose an encoding. We represent a wind farm with a finite number of grid points (where turbines can be installed). In this layout, the $1s$ represent a turbine being present in a specific grid point, and $0s$ represent an empty space in the grid. Note that, the Gaussian process as the Bayesian model is trained on two dimensional real coordinates, which are then encoded to 0 and 1 to maximise the acquisition function. An illustration of encoding representation of a wind farm (with 20×20 grid points) is shown in Figure 2.

4 RESULTS AND DISCUSSION

We applied the multi-objective BO over sets on a wind farm with maximum of 400 turbines spaced in 20×20 grid[1]. The rest of experimental settings are mentioned in the supplementary material. The joint distribution of wind speed and direction estimated with kernel density estimation is shown in Figure 2.

[1]We used the EHVI implementation available at https://liacs.leidenuniv.nl/~csmoda/index.php?page=code

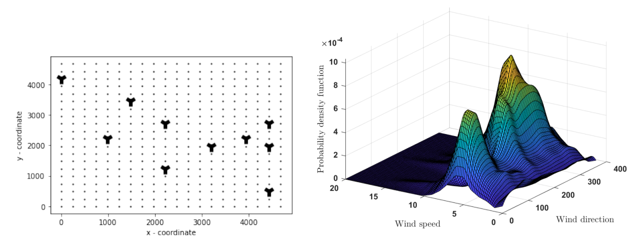

Figure 2: Visual representation of the binary encoding (left). The $1s$ are the location of turbines and $0s$ are the empty space. Probability density function of the joint distribution of wind speed (m/sec) and wind direction (right).

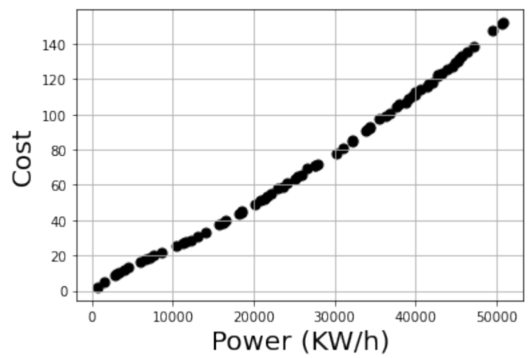

Figure 3: Approximated Pareto front - Power is to be maximised and cost is to minimised.

After running the algorithm for 100 expensive evaluations, we obtained an approximated Pareto front and is shown in Figure 3 (for the run with median hypervolume value). Each point on the approximated Pareto front represents a wind farm with a different number of turbines.

The set based kernel in Gaussian process consider the correlation between wind farms. In other words, two wind farms with similar x and y coordinates have a large correlation compared to two wind farms with different x and y coordinates. For instance, consider six wind farms (a, b, c, d, e, f) shown in Figure 4. These wind farms have different number of turbines. Some of the wind farms e.g. ('a' and 'b'), ('c' and 'd') and ('e' and 'f') have turbines in the similar locations. Therefore, we expect to see a higher correlation between these pairs compared to e.g. ('a' and 'e') or ('b' and 'd'). We show the covariance matrix representing the correlation between these wind farms in the figure. As can been seen, obviously the correlation values are the largest among diagonal elements. After the diagonal elements, the correlation values follow the kernel calculations. For example, the wind farms in the decreasing order of correlation with wind farm 'a' are ('a', 'b', 'd', 'e', 'c', 'f') and with wind farm 'e' are ('e', 'f', 'a', 'c', 'd', 'b').

5 CONCLUSIONS

In this work, we applied Multi-objective BO with expected hypervolume improvement (EHVI) as the acquisition function to handle

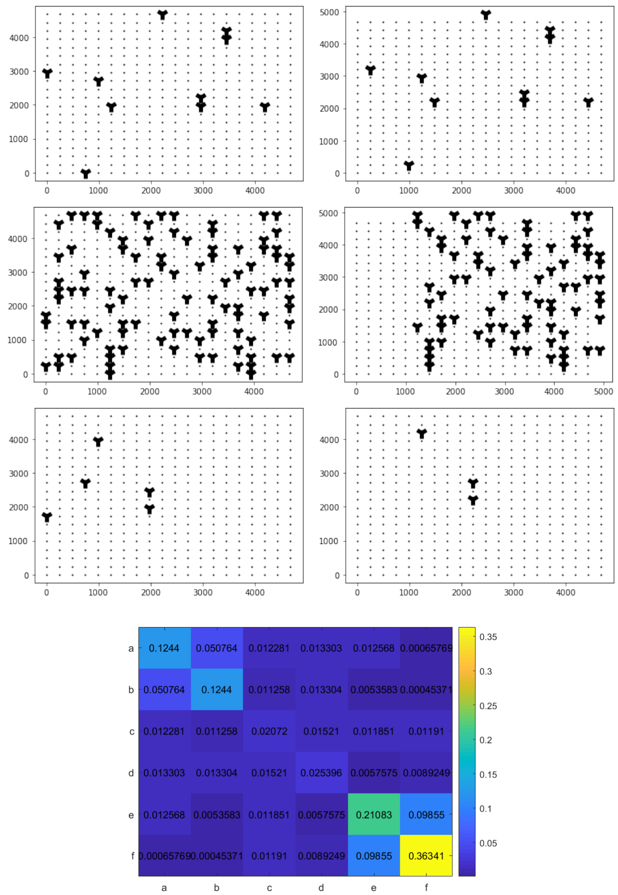

Figure 4: Six wind farms and correlation among them. The correlation matrix is generated with length scale, $l = (82 \times 3)$, amplitude $\sigma^2 = 1$, and noise variance $\sigma_n^2 = 1$.

the computationally expensive optimisation problem. To handle the uncertainty in wind direction and wind speed, we modelled the data with kernel density estimation. As the power output of the wind farm relied on design sets, we utilised the kernel over sets in Gaussian process and embedded it into the multi-objective BO. The method was able to handle and quantify the correlation between different wind farms with different number of turbines and resulted in a set of approximated Pareto optimal solutions.

This paper presented the first attempt at trying to solve the wind farm layout optimisation using Multi-objective BO over sets. We believe that there are particular aspects of the model that can be further investigated and improved. For example, finding correlations between wind farms could be improved with a different set based kernel. Moreover, the computational complexity of set based kernel is $O(N^2|X|^2d)$, which is higher than the computational complexity $O(N^3)$ of traditional Gaussian process. Further research could provide ways of reducing the computational cost of the model. In this work, we built a surrogate model for the cost function which was not computationally expensive and was a function of the number of turbines. This modelling of the cost function may not be required to maximise the EHVI. Working on such heterogeneous objective functions in Bayesian optimisation is also a future research direction. Utilising different wake models including computational fluid dynamic solvers will also be considered as one of the main future research directions. We believe that this work can lead to an increase interest in set based optimisation especially in Multi-objective BO which will widen its application to other research areas, especially where the objective functions rely on a set of solutions.

ACKNOWLEDGMENTS

The authors would like to thank Prof. K. Mitra, Prof. J. Fieldsend, Prof. R. Everson, Dr S.S. Miriyala and Dr P.D. Pantula for initial discussions on optimising wind farms. This research was partially supported by South Asia Development Fund at the University of Exeter, UK.

REFERENCES

[1] [n. d.]. Paris Agreement. https://ec.europa.eu/clima/policies/international/negotiations/paris_en, 2019.
[2] T. Chugh. 2020. Scalarizing Functions in Bayesian Multiobjective Optimization. In *2020 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 1–8.
[3] T. Chugh and M. López-Ibáñez. 2021. Maximising Hypervolume and Minimising ϵ-Indicators using Bayesian Optimisation over Sets. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*. ACM, 1326–1334.
[4] M. Emmerich, A.H. Deutz, and J.W. Klinkenberg. 2011. Hypervolume-based expected improvement: Monotonicity properties and exact computation. In *Proceedings of the IEEE Congress on Evolutionary Computation*. IEEE, 2147–2154.
[5] M. Emmerich, K. Giannakoglou, and B. Naujoks. 2006. Single- and multiobjective evolutionary optimization assisted by Gaussian random field metamodels. *IEEE Transactions on Evolutionary Computation* 10 (2006), 421–439.
[6] EnigieOPENdata. [n. d.]. La Haute Borne Data. https://opendata-renewables.engie.com/explore/
[7] A. Evans, V. Strezov, and T.J. Evans. 2009. Assessment of sustainability indicators for renewable energy technologies. *Renewable and Sustainable Energy Reviews* 13, 5 (2009), 1082–1088.
[8] J. Kim, M. McCourt, T. You, S. Kim, and S. Choi. 2021. Bayesian optimization with approximate set kernels. *Machine Learning* 110, 5 (2021), 857–879.
[9] A. Kusiak, H. Zheng, and Z. Song. 2009. On-line monitoring of power curves. *Renewable Energy* 34, 6 (2009), 1487–1493.
[10] P. Mittal and K. Mitra. 2019. Robust Wind Farm Layout Optimization under Uncertainty. In *2019 Sixth Indian Control Conference (ICC)*. IEEE, 164–169.
[11] P. Mittal and K. Mitra. 2020. In search of flexible and robust wind farm layouts considering wind state uncertainty. *Journal of Cleaner Production* 248 (2020), 119195.
[12] G. Mosetti, C. Poloni, and B. Diviacco. 1994. Optimization of wind turbine positioning in large windfarms by means of a genetic algorithm. *Journal of Wind Engineering and Industrial Aerodynamics* 51, 1 (1994), 105–116.
[13] C. E. Rasmussen and C. K. I. Williams. 2006. *Gaussian processes for machine learning*. The MIT Press.
[14] R. Shakoor, M. Y. Hassan, A. Raheem, and Y.-K. Wu. 2016. Wake effect modeling: A review of wind farm layout optimization using Jensens model. *Renewable and Sustainable Energy Reviews* 58 (2016), 1048–1059.