Bayesian Optimisation Against Climate Change: Applications and Benchmarks

Sigrid Passano Hellan [1] **Christopher G Lucas** [1] **Nigel H Goddard** [1]

Abstract

Bayesian optimisation is a powerful method for optimising black-box functions, popular in settings where the true function is expensive to evaluate and no gradient information is available. Bayesian optimisation can improve responses to many optimisation problems within climate change for which simulator models are unavailable or expensive to sample from. While there have been several feasibility demonstrations of Bayesian optimisation in climate-related applications, there has been no unifying review of applications and benchmarks. We provide such a review here, to encourage the use of Bayesian optimisation in important and well-suited application domains. We identify four main application domains: material discovery, wind farm layout, optimal renewable control and environmental monitoring. Our contributions are: a) identifying a representative range of benchmarks, providing example code where necessary; b) introducing a new benchmark, LAQN-BO; and c) promoting a wider use of climate change applications among Bayesian optimisation practitioners.

1. Introduction

The use of machine learning (ML) to tackle climate change is gaining traction, including ML-wide surveys of relevant applications (Rolnick et al., 2022; Donti & Kolter, 2021). To facilitate technical innovation and its adoption into practice, it is important to complement these reviews with concise summaries of problems that specific ML frameworks are well-matched to, along with example data sets. Here, we provide a summary, paired with data sets, for Bayesian optimisation. This is important for three reasons: a) it reduces start-up costs and overhead, as the time-consuming task of identifying suitable data sets does not have to be repeated; b) it helps researchers focus on problems that are timely and/or important, and have a structure that is suitable for the framework(s) they are using; and c) it makes it easier to compare methods as numerical results can be compared directly. The benchmarks should be similar to real-world applications, because that provides more realistic expectations of performance. It also shows which versions of Bayesian optimisation are most likely to perform well, for instance in terms of choice of acquisition function and underlying model. More realistic benchmarks also let us identify challenges which still need to be solved, for instance in terms of generating useful priors from related data.

Bayesian optimisation (BO), is a machine-learning approach to black-box optimisation (Shahriari et al., 2015). It is particularly suited to problems that:

a. have a complex unknown structure which can only be efficiently modelled through a surrogate model, e.g. the conductivity of solar panel materials;
b. are expensive or slow to evaluate, requiring sample efficiency from trying out fewer bad or uninformative choices, e.g. when discovering new materials that must be synthesised in order to be evaluated;
c. require identifying extrema, either for optimisation, e.g. maximising power generation, or for intervention, e.g. emission monitoring.

An additional benefit of Bayesian optimisation is that it explicitly models its uncertainty, which can be used to determine when to stop optimising (Makarova et al., 2022), or to trade off exploration and exploitation (De Ath et al., 2021).

Bayesian optimisation consists of fitting a surrogate model, typically a Gaussian process (Williams & Rasmussen, 2006), to observed data, and using the surrogate model to choose what input value to evaluate next. Then the surrogate model is adjusted, another input value evaluated, and so on. To choose the next input value an acquisition function is used, which balances exploration and exploitation. As an illustrative example, visualised in Figure 1, Bayesian optimisation can be used to find the angle of a solar panel that maximises power output: (1) model the output as a function of angle

[1] School of Informatics, University of Edinburgh, UK. Correspondence to: Sigrid Passano Hellan <s.p.hellan@ed.ac.uk>.

Presented at the Data-centric Machine Learning Research (DMLR) Workshop at the 40^{th} International Conference on Machine Learning, Honolulu, Hawaii, USA. 2023.

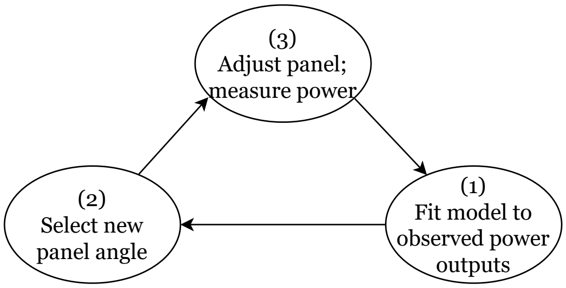

Figure 1: Application of Bayesian optimisation to optimising the power generation of a solar panel by adjusting the panel angle.

based on all available data and/or domain knowledge; (2) select a promising angle to try next, e.g., an angle that is a good candidate for maximising power output or that is maximally informative about the best angle; and (3) try this angle before updating the model and returning to Step 1.

We present a summary of Bayesian optimisation as applied to climate change problems. We identify the main application domains, emphasise how they impact climate change responses, and identify beginner-friendly benchmarks for each. We found a benchmark to be lacking for environmental monitoring, and therefore provide one, the LAQN-BO benchmark, based on air pollution data (Imperial College London, 1993). The four use cases are: *material discovery* — accelerating the development of e.g. new solar panels; *wind farm layouts* — choosing where to place individual turbines; *optimal renewable control* — choosing operating parameters of solar panels or wind turbines; and *environmental monitoring* — choosing where to place sensors. (Bliek, 2022) is a related survey of Bayesian optimisation and other surrogate-based optimisation methods. Where Bliek takes a higher-level approach, focusing on classifying previous work by methodology, we instead group work by applications in order to identify data sets and applications for new projects. When selecting benchmarks, we have emphasised ease of adoption. So we have only considered ones that are publicly available, and well documented.

This paper is predominantly concerned with Bayesian optimisation. However, Bayesian optimisation is closely related to other methods, like Bayesian experimental design (Rainforth et al., 2023), and we have included some examples of applications of these methods as well, e.g. (Tran & Ulissi, 2018; Kleinegesse & Gutmann, 2020). Bayesian optimisation can be used within Bayesian experimental design (Valentin et al., 2023), and can also be seen as a special case of Bayesian experimental design (Rainforth et al., 2023), especially when using information-theoretic acquisition functions. Both are powerful methods, with useful applications in the response to climate change and similar challenges, such as the choice of priors and scaling to large numbers of samples. Our focus on Bayesian optimisation does not reflect a preference of one method over the other, but rather that more work has been done applying Bayesian optimisation against climate change than for Bayesian experimental design. Because the methods are so closely related, and because our focus is on the applications and benchmarks, not on the details of the methods used for solving them, we have opted to include some examples of related methods as well.

Bayesian optimisation is not a complete solution, but an important tool when combined with domain knowledge, and can help improve responses to climate change. Performance in real applications is the ultimate test of any ML method, including Bayesian optimisation. Simulated problems often fail to anticipate crucial challenges, e.g. in terms of finding suitable inductive biases (Hellan et al., 2022). Currently, Bayesian optimisation is often evaluated on synthetic functions or hyperparameter optimisation (HPO) benchmarks, either exclusively (AV et al., 2022; Papenmeier et al., 2022; Hvarfner et al., 2022; Thebelt et al., 2022) or mostly (Oliveira et al., 2022; Folch et al., 2022; Song et al., 2022). But HPO is just one kind of problem. There are other less used benchmarks, like robotics (Song et al., 2022; Nguyen et al., 2022) and chemistry simulations (Folch et al., 2022; Tu et al., 2022). (Ramesh et al., 2022) test their method on optimising the control of an airborne wind energy system, but has not made their setup available as a benchmark. To increase the impact of Bayesian optimisation, and its prevalence, we should adopt a broader set of benchmarks. That would not only demonstrate which methods work well in practice, but open up new research problems to solve. Using benchmarks from climate change has the advantage of doing so in a dynamic field, increasing the chance of adoption, and with real impact. Adopting climate change motivated benchmarks, and using them to compare methods, will make Bayesian optimisation more relevant outside academia. We now proceed to the main applications domains and their respective benchmarks, starting with material discovery.

2. Material discovery

Material discovery means developing materials with superior combinations of properties (Frazier & Wang, 2016), e.g. more efficient solar cells. Bayesian optimisation and related methods have been suggested for optimising a wide range of materials, from concrete (Severson et al., 2021) and Direct Air Capture of CO_2 (Ortiz-Montalvo et al., 2021) to solar panel glass (Haghanifar et al., 2021), electrifying the transport and chemical industry sectors (Annevelink et al., 2022), and for electrocatalysts for CO_2 reduction and H_2 production (Tran & Ulissi, 2018; Zhong et al., 2020; Frey et al., 2022). The benefits are in reducing greenhouse gas emissions or increasing their capture; making renewable energy generation or storage more efficient; or making more sectors able to use renewable energy. If evaluating the material requires synthesising and testing it, there are huge potential gains in time and cost by only doing so on promising candidates identified by Bayesian optimisation.

(Liang et al., 2021) provide three material discovery data sets (Sun et al., 2021; Bash et al., 2021; Mekki-Berrada et al., 2021) related to solar panels, with 3-5 features each and between 94 and 178 unique evaluated data points. Using Figure 1 to illustrate the Bayesian optimisation process, the steps become: (1) model the material performance as a function of ingredient proportions used; (2) select a new combination of ingredient proportions to try; and (3) produce and evaluate the material. The data sets enable us to replace (3) with looking up the evaluations in a table.

- **Benchmark:** `https://github.com/PV-Lab/Benchmarking` (Liang et al., 2021)
- **Features:** Material properties: 3-5 dimensions
- **Data type:** CSV files
- **Sampling:** Manufacture material. In benchmark replaced by table look-up
- **Objective:** Material performance: conductivity, absorbance spectrum score or stability
- **Impact:** Better solar panels, which gives more renewable generation and lower climate gas emissions

3. Wind farm layout

Wind turbines are typically grouped together in wind farms, with several turbines in relatively close proximity, to reduce installation costs and environmental impacts. Before construction, the locations of individual wind turbines must be planned. Wind turbines work by extracting kinetic energy from the air, so the wind is weaker and more turbulent after flowing past a turbine. This leads to less power generation (Park & Law, 2015) and greater dynamic loads and fatigue on the downwind turbines (Dong et al., 2022). This is illustrated in Figure 2, where the two turbines need to be placed within the green area. Determining the optimal wind turbine layout is a difficult but important optimisation problem, as more renewable electricity can be generated without requiring more turbines to be installed. The efficiency of Bayesian optimisation is important for this problem, as the type of simulations used can take 15 seconds to run on one CPU even for five wind turbines (Bliek et al., 2021). Bayesian optimisation can be applied to the problem by replacing the steps in Figure 1 with (1) modelling the collective power output as a function of the wind turbine placements; (2) selecting a new combination of locations to evaluate; and (3) running the power output simulation.

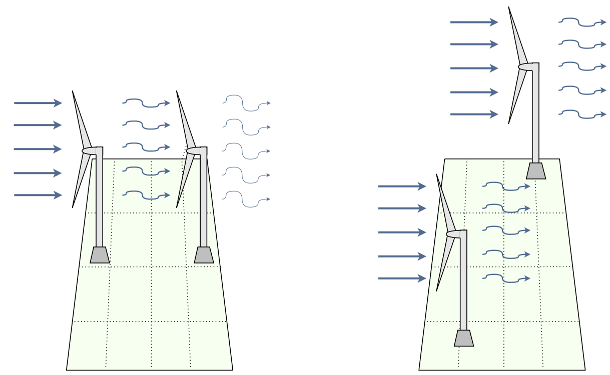

Figure 2: Planning of wind turbines within limits shown in green. The wind is shown by the blue arrows: wider arrows indicate stronger wind and more undulating arrows indicate more turbulence. In the layout on the left one turbine is directly downwind of the other, resulting in weaker and more turbulent wind. To the right, both turbines get unhindered wind.

Bliek et al.'s (2021) set of benchmarking problems and baselines include a wind farm layout problem, *windwake*, based on FLORIS (NREL, 2020) simulations. The benchmark assumes a fixed number of wind turbines; this is relaxed in (Chugh & Ymeraj, 2022), where the two objectives of cost and power generation are jointly optimised, using simulations based on (Pedersen et al., 2019). In related pieces of work, (Mern et al., 2021) build wind maps for later layout planning, and (Tillmann et al., 2020) plan the layout of bifacial solar panel arrays. We use the benchmark in (Bliek et al., 2021) as it has been been prepared for further use.

- **Benchmark:** `https://github.com/AlgTUDelft/ExpensiveOptimBenchmark` (Bliek et al., 2021)
- **Features:** Spatial locations: 10 dimensions. 2 for each of 5 turbines
- **Data type:** Simulation
- **Sampling:** Simulate layout
- **Objective:** Energy production
- **Impact:** More renewable generation, which reduces climate gas emissions

4. Optimal renewable control

Having planned the layout of a new renewable power generation plant, the next question is how best to operate it. For wind turbines the yaw and pitch angles need to be set (Doekemeijer et al., 2019; Park & Law, 2015; Park, 2020; Yang et al., 2022), as well as the induction factor (Park, 2020) and electric load resistance (Park & Law, 2015). For solar panels the voltage level applied (Abdelrahman et al., 2016; Lyden et al., 2018) and the angle of the panel (Abel et al., 2018) can be adjusted. By adjusting these the amount of energy generated can be maximised, which in turn reduces the amount of fossil fuel consumed for a given amount of electricity demand. A review of different control methodologies for wind turbines is given in (Dong et al., 2022). For wind farms, the optimisation should occur jointly, as the available wind energy for a downstream turbine is impacted by the operating conditions of the upstream turbine (Park & Law, 2015). Bayesian optimisation is suitable for this problem as it allows the response of the non-linear problem to be learnt through data collection. Extensive work on wind turbine control has also been done using Bayesian Ascent, a version of Bayesian optimisation which limits the changes in inputs between iterations (Park & Law, 2016a;b; Park et al., 2016; 2017; 2018). Relatedly, (Moustakis et al., 2019) use Bayesian optimisation to tune the parameters of traditional control methods, and (Fiducioso et al., 2019) consider the problem of controlling a HVAC system. In (Mulders et al., 2020), Bayesian optimisation is used to reduce blade fatigue. (Andersson & Imsland, 2020) use a method related to Bayesian optimisation for wind turbine control.

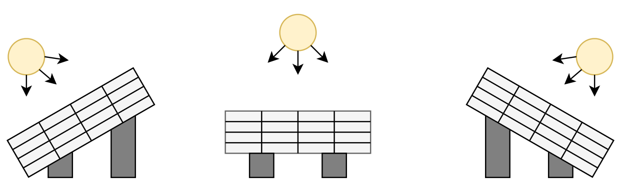

Figure 3: Adjusting solar panels to optimise power generation as the sun moves across the sky.

(Abel et al., 2018) contribute a benchmark for solar panel optimisation[1]. The goal is to maximise the energy yield by adjusting the angles of the solar panel, as illustrated in Figure 3. They evaluate contextual bandits and reinforcement learning methods, but the problem is also suitable for Bayesian optimisation. We provide a basic example of this at `https://github.com/sighellan/solar_panels_rl`[2]. Figure 1 shows how Bayesian optimisation is applied to the problem. We model power output as a function of panel angles; decide on a new angle to try; and run the simulation to evaluate the new angle. The challenge is not to locate the sun; its location can be calculated and is given as input to the methods. Instead, the idea is to adapt to other changes like cloud cover, shading from nearby buildings and changes in the amount of reflected sunlight, for instance after snowfall. However, many solar panels are not designed to track the sun, and instead keep a fixed angle. Wind turbines have more control settings, interact in less predictable ways, and have seen more applications of Bayesian optimisation (see previous paragraph). We chose the benchmark by (Abel et al., 2018) because it is well-documented and publicly available. But real-world benefits are more likely to come from applications to wind turbines than to solar panels.

- **Benchmark:** `https://github.com/david-abel/solar_panels_rl` (Abel et al., 2018)
- **Features:** Panel direction: 2 dimensions
- **Data type:** Simulation
- **Sampling:** Simulate collected sunlight
- **Objective:** Collected energy
- **Impact:** More renewable generation, which reduces climate gas emissions

5. Environmental monitoring

An important problem within environmental monitoring is that of choosing where to place the monitoring sensors. This is difficult as the planning has to be done without knowing the usefulness of each site, as that depends on future data. Bayesian optimisation and related methods are useful tools in this setting, as they build probabilistic models of the environmental characteristic of interest — e.g. air pollution — which they use to make efficient choices. If we are mainly interested in identifying the worst-hit locations, e.g. for targeting interventions or evaluating compliance with legal limits, Bayesian optimisation lets us do that efficiently. To build accurate environmental maps for an entire area we can instead use Bayesian experimental design, for instance if we want to know the water quality throughout a lake.

Monitoring is important for tracking the impacts of climate change, including oceans (Sanchez-Pi et al., 2020), for locating gas leakages (Asenov et al., 2019; Gao & Bai, 2022; Kleinegesse & Gutmann, 2020), and the related task of air pollution monitoring (Ainslie et al., 2009; Morere et al., 2017; Marchant & Ramos, 2012; Singh et al., 2010; Hellan et al., 2020; 2022; Leu & Ho, 2020). (Samaniego et al., 2021) provide a benchmark for path-planning based on water quality monitoring, also used by (Folch et al., 2022), but the benchmark uses synthetic data.

The LAQN-BO benchmark

We present a new benchmark for the climate change Bayesian optimisation community, using air pollution data from the London Air Quality Network (LAQN) (Imperial College London, 1993). Air pollution monitoring is related to climate change in that the sources and hence required interventions are often interlinked, e.g. car exhaust. Also, environmental changes are often sped up by climate change, making their monitoring more important. The optimisa-

[1]To make it work, we used v0.784 of the simple_rl Python package (Abel, 2019).

[2]We built the example using BoTorch (Balandat et al., 2020). In addition to integrating our basic Bayesian optimisation implementation into the framework, we also provide a requirements file and port the code to Python 3.

tion problems in the LAQN-BO benchmark are similar to those in (Hellan et al., 2022). But while that focuses on Bayesian optimisation methodology, we focus on the problems themselves. We provide simple scripts and step-by-step instructions for generating the problems and put them in the context of other climate-related Bayesian optimisation problems and benchmarks.

The optimisation objective is the NO_2 concentration, i.e. to find the location of maximum pollution from the set of available locations. This is useful as higher pollution concentrations mean greater impacts on human health. Again using Figure 1 as a basis, we replace the steps with: (1) model the NO_2 concentration as a function of spatial location; (2) determine a new location to evaluate; and (3) measure the pollution in the new location. We use the provided benchmark to replace (3) with looking up the historical measurement. We construct a training set from the 2015 data, and a test set from the 2016 data. Each problem corresponds to data from a single day, so multiple days give multiple problems. As in (Hellan et al., 2022), we filter out days when less than 40 stations collected measurements, and only use the 'Roadside' sensors, resulting in 214 training problems and 365 test problems. We log-transformed the data, and standardised it using the mean and standard deviation from the training set. Example problems are shown in Figure 4.

The LAQN-BO benchmark has several advantages. Firstly, it uses real data. Secondly, we present a *set* of related problems, so that methods can be trained on the training set, and evaluated on the test set. Finally, the provided scripts can be easily adapted to generate new Bayesian optimisation problems using other pollutants or other years of the extensive available LAQN data. None of the other benchmarks combine all of these advantages. A disadvantage is that data is not available everywhere, only at locations with existing sensors, limiting the available search space. This corresponds to having to choose from a sparse set of locations. For instance, if we are mounting sensors to lampposts, we can only choose between the locations with lampposts.

- **Benchmark:** `https://github.com/sighellan/LAQN-BO`
- **Features:** Spatial coordinates: 2 dimensions
- **Data type:** Python class
- **Sampling:** Place pollution sensor. In benchmark replaced by table look-up
- **Objective:** NO_2 concentration
- **Impact:** Better monitoring of air pollution, enabling targeted interventions

6. Benchmark comparison

The four benchmarks identified in our survey have different characteristics, as is summarised in Table 1. They represent different key challenges for Bayesian optimisation deployments, and provide breadth for evaluating different types of Bayesian optimisation methods. The key challenge of the materials benchmark is its small size, meaning there is little opportunity to fine-tune the methods. For LAQN-BO it is that of constructing priors from the training data, as the problems have as little as 40 evaluations each, requiring very sample-efficient learning. The wind farm layout benchmark also requires efficient use of data. But in contrast to LAQN-BO, it does not provide training data for learning priors. The key challenge for the renewable control benchmark is dealing with large numbers of samples, as the panel direction is optimised regularly over many days. It is therefore a good candidate for testing Bayesian optimisation built on scalable Gaussian processes (Liu et al., 2020). It also requires keeping track of context, e.g. time, as the sun moves across the sky. An additional contrast to the other benchmarks is that the performance throughout the optimisation matters, as it directly impacts the amount of electricity generated.

The materials benchmark is the easiest to start using, as the data is provided in CSV files, followed by that for environmental monitoring, LAQN-BO, which provides a simple Python interface to the data. The benchmarks for wind farm layout and renewable control require more setup, as methods need to be interfaced with the simulators. Their advantage is that they are easier to extend, and can be evaluated for more input values. LAQN-BO can be easily extended to more problems using the LAQN data (Imperial College London, 1993), but extensions to other base data sets would require more work. The materials benchmark is the hardest to extend, as it relies on results from synthesising and evaluating materials.

7. Conclusion and other applications

Bayesian optimisation and related methods have many more climate-related applications: learning energy consumptions of individual household appliances (Jia et al., 2019); optimising charging protocols for electric vehicle batteries (Attia et al., 2020); scheduling smart appliances to smooth out demand curves (Tabakhi et al., 2020); determining vibration suppression parameters for floating wind turbines (Zhang et al., 2022); producing policies to reduce the impact of livestock diseases made more prevalent by global warming (Spooner et al., 2020); and tuning parameters for HVAC system and building models to reduce cooling and heating energy demands (Zhan et al., 2022; Chakrabarty et al., 2021a;b). Bayesian optimisation also has applications in improving climate models, by targeting informative training data (Watson-Parris, 2021). And it has been used

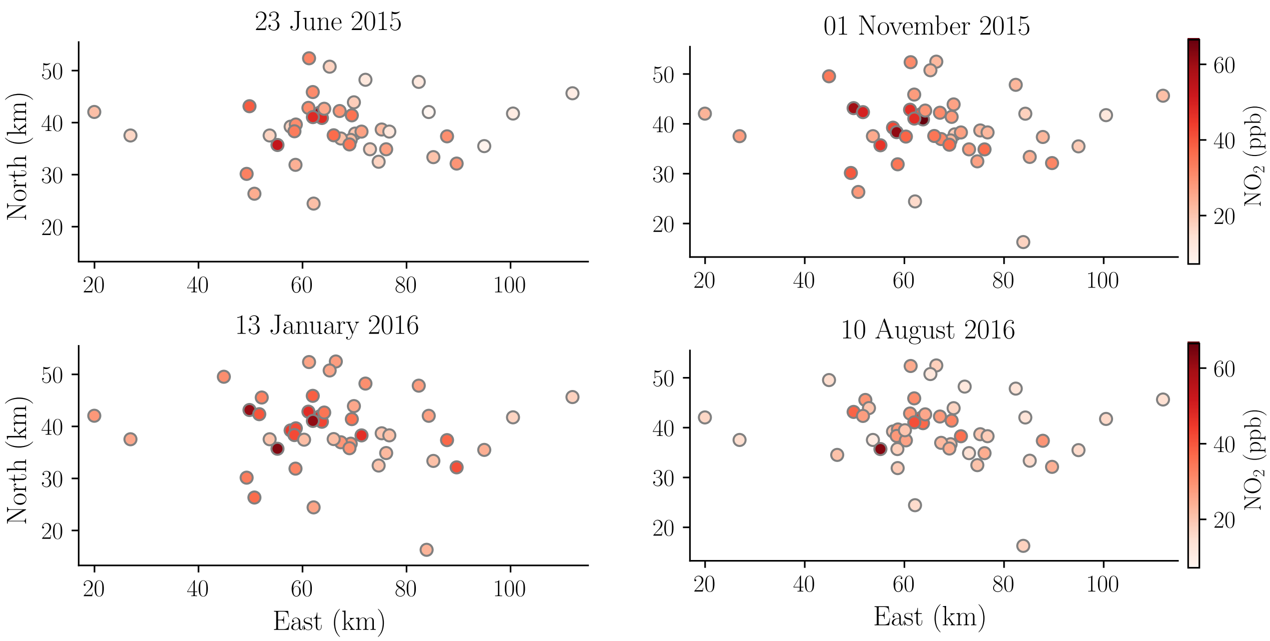

Figure 4: Example LAQN-BO problems from the training set (2015, top) and test set (2016, bottom). Not all sensors are available on every day. The overall pollution level varies, compare the top left and right plots. The locations of maxima also varies, compare the right top and bottom plots. Most of the sensors are clustered together in central London.

Table 1: Comparison of benchmarks.

Benchmark	Type	Key challenge	Dimensionality	# problems
Materials (Liang et al., 2021)	Real	Small size	3/5	3
Wind farm layout (Bliek et al., 2021)	Simulation	Sample efficiency	10	1
Renewable control (Abel et al., 2018)	Simulation	Scaling, context	2	10
Environmental monitoring (LAQN-BO)	Real	Priors	2	214+365

for parameter tuning of a global land surface model (Druel et al., 2017), a landslide model (Pradhan et al., 2021) and forecasting models for electricity generation and demand (Trivedi & Khadem, 2021).

To encourage more Bayesian optimisation practitioners to work on climate change applications, an important next step is to establish more benchmarks, standard data sets or simulators, as we do with the LAQN-BO benchmark. More challenging benchmarks should also be established, e.g. for constrained or multi-objective optimisation, to encompass the full complexity of real problems. For example: outside a simulation, some of the ground might be unsuitable for building wind turbines; and when building renewable energy generators there are multiple competing objectives (Wu et al., 2018; Flecker et al., 2022). Additionally, there might be constraints for controlling the generators, due to increased energy costs and fatigue when changing settings. The energy cost of changing the panel angles is included in the renewable control benchmark (Abel et al., 2018), so Bayesian Ascent (Park & Law, 2016a) should be tested, as it limits the difference in input — panel angles — between iterations. As a further extension, the layout and control of wind farms can be optimised jointly (Chen et al., 2022).

In the future, all these benchmarks should be brought together and their interfaces aligned, to create a standardised, climate-themed benchmark suite. By identifying four benchmarks representing the main application domains, introducing a new benchmark, and highlighting the potential for Bayesian optimisation, we take an important step in facilitating future research.

Acknowledgements

We would like to thank Linus Ericsson, Simon Valentin and Michael Gutmann at the University of Edinburgh for helpful suggestions for improving the paper.

This work was supported by the EPSRC Centre for Doctoral Training in Data Science, funded by the UK Engineering and Physical Sciences Research Council (grant EP/L016427/1) and the University of Edinburgh.

References

Abdelrahman, H., Berkenkamp, F., Poland, J., and Krause, A. Bayesian optimization for maximum power point tracking in photovoltaic power plants. In *2016 European Control Conference (ECC)*, pp. 2078–2083. IEEE, 2016.

Abel, D. simple_rl: Reproducible reinforcement learning in python. 2019.

Abel, D., Williams, E. C., Brawner, S., Reif, E., and Littman, M. L. Bandit-based solar panel control. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

Ainslie, B., Reuten, C., Steyn, D. G., Le, N. D., and Zidek, J. V. Application of an entropy-based Bayesian optimization technique to the redesign of an existing monitoring network for single air pollutants. *Journal of Environmental Management*, 90(8):2715–2729, June 2009. ISSN 0301-4797. doi: 10.1016/j.jenvman.2009.02.016. URL https://www.sciencedirect.com/science/article/pii/S0301479709000711.

Andersson, L. E. and Imsland, L. Real-time optimization of wind farms using modifier adaptation and machine learning. *Wind Energy Science*, 5(3):885–896, 2020.

Annevelink, E., Kurchin, R., Muckley, E., Kavalsky, L., Hegde, V. I., Sulzer, V., Zhu, S., Pu, J., Farina, D., Johnson, M., et al. Automat: Automated materials discovery for electrochemical systems. *MRS Bulletin*, pp. 1–9, 2022.

Asenov, M., Rutkauskas, M., Reid, D., Subr, K., and Ramamoorthy, S. Active Localization of Gas Leaks Using Fluid Simulation. *IEEE Robotics and Automation Letters*, 4(2):1776–1783, April 2019. ISSN 2377-3766. doi: 10.1109/LRA.2019.2895820.

Attia, P. M., Grover, A., Jin, N., Severson, K. A., Markov, T. M., Liao, Y.-H., Chen, M. H., Cheong, B., Perkins, N., Yang, Z., et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature*, 578(7795):397–402, 2020.

AV, A. K., Rana, S., Shilton, A., and Venkatesh, S. Human-ai collaborative Bayesian optimisation. *Advances in Neural Information Processing Systems*, 35:16233–16245, 2022.

Balandat, M., Karrer, B., Jiang, D. R., Daulton, S., Letham, B., Wilson, A. G., and Bakshy, E. BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization. In *Advances in Neural Information Processing Systems 33*, 2020. URL http://arxiv.org/abs/1910.06403.

Bash, D., Cai, Y., Chellappan, V., Wong, S. L., Yang, X., Kumar, P., Tan, J. D., Abutaha, A., Cheng, J. J., Lim, Y.-F., et al. Multi-fidelity high-throughput optimization of electrical conductivity in P3HT-CNT composites. *Advanced Functional Materials*, pp. 2102606, 2021.

Bliek, L. A survey on sustainable surrogate-based optimisation. *Sustainability*, 14(7):3867, 2022.

Bliek, L., Guijt, A., Karlsson, R., Verwer, S., and de Weerdt, M. Expobench: Benchmarking surrogate-based optimisation algorithms on expensive black-box functions. *arXiv preprint arXiv:2106.04618*, 2021.

Chakrabarty, A., Danielson, C., Bortoff, S. A., and Laughman, C. R. Accelerating self-optimization control of refrigerant cycles with Bayesian optimization and adaptive moment estimation. *Applied Thermal Engineering*, 197:117335, 2021a.

Chakrabarty, A., Maddalena, E., Qiao, H., and Laughman, C. Scalable Bayesian optimization for model calibration: Case study on coupled building and HVAC dynamics. *Energy and Buildings*, 253:111460, 2021b.

Chen, K., Lin, J., Qiu, Y., Liu, F., and Song, Y. Joint optimization of wind farm layout considering optimal control. *Renewable Energy*, 182:787–796, 2022.

Chugh, T. and Ymeraj, E. Wind farm layout optimisation using set based multi-objective Bayesian optimisation. In *Genetic and Evolutionary Computation Conference Companion (GECCO '22 Companion)*, pp. 695–698. ACM, New York, NY, USA, 07 2022. URL https://doi.org/10.1145/3520304.3528951.

De Ath, G., Everson, R. M., Rahat, A. A., and Fieldsend, J. E. Greed is good: Exploration and exploitation trade-offs in Bayesian optimisation. *ACM Transactions on Evolutionary Learning and Optimization*, 1(1):1–22, 2021.

Doekemeijer, B. M., Van Der Hoek, D. C., and van Wingerden, J.-W. Model-based closed-loop wind farm control for power maximization using Bayesian optimization: a large eddy simulation study. In *2019 IEEE Conference on Control Technology and Applications (CCTA)*, pp. 284–289. IEEE, 2019.

Dong, H., Xie, J., and Zhao, X. Wind farm control technologies: From classical control to reinforcement learning. *Progress in Energy*, 2022.

Donti, P. L. and Kolter, J. Z. Machine learning for sustainable energy systems. *Annual Review of Environment and Resources*, 46:719–747, 2021.

Druel, A., Peylin, P., Krinner, G., Ciais, P., Viovy, N., Peregon, A., Bastrikov, V., Kosykh, N., and Mironycheva-Tokareva, N. Towards a more detailed representation of high-latitude vegetation in the global land surface model ORCHIDEE (ORC-HL-VEGv1. 0). *Geoscientific Model Development*, 10(12):4693–4722, 2017.

Fiducioso, M., Curi, S., Schumacher, B., Gwerder, M., and Krause, A. Safe contextual Bayesian optimization for sustainable room temperature PID control tuning. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, pp. 5850–5856. International Joint Conferences on Artificial Intelligence Organization, 7 2019. doi: 10.24963/ijcai.2019/811. URL https://doi.org/10.24963/ijcai.2019/811.

Flecker, A. S., Shi, Q., Almeida, R. M., Angarita, H., Gomes-Selman, J. M., García-Villacorta, R., Sethi, S. A., Thomas, S. A., Poff, N. L., Forsberg, B. R., et al. Reducing adverse impacts of amazon hydropower expansion. *Science*, 375(6582):753–760, 2022.

Folch, J. P., Zhang, S., Lee, R., Shafei, B., Walz, D., Tsay, C., van der Wilk, M., and Misener, R. SnAKe: Bayesian optimization with pathwise exploration. *Advances in Neural Information Processing Systems*, 35:35226–35239, 2022.

Frazier, P. I. and Wang, J. Bayesian optimization for materials design. In *Information science for materials discovery and design*, pp. 45–75. Springer, 2016.

Frey, D., Neyerlin, K., and Modestino, M. A. Bayesian optimization of electrochemical devices for electrons-to-molecules conversions: The case of pulsed CO2 electroreduction. *ChemRxiv*, 2022. doi: 10.26434/chemrxiv-2022-n09hc.

Gao, T. and Bai, X. Bayesian optimization-based three-dimensional, time-varying environment monitoring using an UAV. *Journal of Intelligent & Robotic Systems*, 105 (4):1–21, 2022.

Haghanifar, S., Cheng, B., McCourt, M., and Leu, P. Using Bayesian optimization to improve solar panel performance by developing antireflective, superomniphobic glass. 2021. URL https://www.climatechange.ai/papers/icml2021/84.

Hellan, S. P., Lucas, C. G., and Goddard, N. H. Optimising Placement of Pollution Sensors in Windy Environments, December 2020. URL http://arxiv.org/abs/2012.10770. Presented at the NeurIPS 2020 Workshop on AI for Earth Sciences.

Hellan, S. P., Lucas, C. G., and Goddard, N. H. Bayesian optimisation for active monitoring of air pollution. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(11):11908–11916, Jun. 2022. doi: 10.1609/aaai.v36i11.21448. URL https://ojs.aaai.org/index.php/AAAI/article/view/21448.

Hvarfner, C., Hutter, F., and Nardi, L. Joint entropy search for maximally-informed Bayesian optimization. In *Advances in Neural Information Processing Systems*, 2022.

Imperial College London, E. R. G. London Air Quality Network. https://www.londonair.org.uk/LondonAir/Default.aspx, 1993. Accessed: 2021-08-31.

Jia, Y., Batra, N., Wang, H., and Whitehouse, K. Active collaborative sensing for energy breakdown. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pp. 1943–1952, 2019.

Kleinegesse, S. and Gutmann, M. U. Bayesian experimental design for implicit models by mutual information neural estimation. In *International Conference on Machine Learning*, pp. 5316–5326. PMLR, 2020.

Leu, F.-Y. and Ho, J.-S. Air pollution source identification by using neural network with Bayesian optimization. In *Innovative Mobile and Internet Services in Ubiquitous Computing: Proceedings of the 13th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS-2019)*, pp. 514–524. Springer, 2020.

Liang, Q., Gongora, A. E., Ren, Z., Tiihonen, A., Liu, Z., Sun, S., Deneault, J. R., Bash, D., Mekki-Berrada, F., Khan, S. A., et al. Benchmarking the performance of Bayesian optimization across multiple experimental materials science domains. *npj Computational Materials*, 7 (1):1–10, 2021.

Liu, H., Ong, Y.-S., Shen, X., and Cai, J. When Gaussian process meets big data: A review of scalable GPs. *IEEE transactions on neural networks and learning systems*, 31 (11):4405–4423, 2020.

Lyden, S., Olding, W., and Darbari, Z. Bayesian optimization for maximum power point tracking in photovoltaic systems. In *2018 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5. IEEE, 2018.

Makarova, A., Shen, H., Perrone, V., Klein, A., Faddoul, J. B., Krause, A., Seeger, M., and Archambeau, C. Automatic termination for hyperparameter optimization. In *International Conference on Automated Machine Learning*, pp. 7–1. PMLR, 2022.

Marchant, R. and Ramos, F. Bayesian optimisation for Intelligent Environmental Monitoring. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2242–2249. IEEE, October 2012. doi: 10.1109/IROS.2012.6385653.

Mekki-Berrada, F., Ren, Z., Huang, T., Wong, W. K., Zheng, F., Xie, J., Tian, I. P. S., Jayavelu, S., Mahfoud, Z., Bash, D., et al. Two-step machine learning enables optimized nanoparticle synthesis. *npj Computational Materials*, 7 (1):1–10, 2021.

Mern, J., Yildiz, A., Sunberg, Z., Mukerji, T., and Kochenderfer, M. J. Bayesian optimized Monte Carlo planning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 11880–11887, 2021. URL https://ojs.aaai.org/index.php/AAAI/article/view/17411.

Morere, P., Marchant, R., and Ramos, F. Sequential Bayesian optimization as a POMDP for environment monitoring with UAVs. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6381–6388. IEEE, May 2017. doi: 10.1109/ICRA.2017.7989754.

Moustakis, N., Mulders, S. P., Kober, J., and van Wingerden, J.-W. A practical Bayesian optimization approach for the optimal estimation of the rotor effective wind speed. In *2019 American Control Conference (ACC)*, pp. 4179–4185. IEEE, 2019.

Mulders, S., Pamososuryo, A., and Van Wingerden, J. Efficient tuning of individual pitch control: a Bayesian optimization machine learning approach. In *Journal of Physics: Conference Series*, volume 1618, pp. 022039. IOP Publishing, 2020.

Nguyen, Q., Wu, K., Gardner, J., and Garnett, R. Local Bayesian optimization via maximizing probability of descent. *Advances in neural information processing systems*, 35:13190–13202, 2022.

NREL. FLORIS. version 2.1.1, 2020. URL https://github.com/NREL/floris.

Oliveira, R., Tiao, L., and Ramos, F. T. Batch Bayesian optimisation via density-ratio estimation with guarantees. *Advances in Neural Information Processing Systems*, 35: 29816–29829, 2022.

Ortiz-Montalvo, D. L., Kusne, A. G., McDannald, A., Siderius, D., Choudhary, K., and Yildirim, T. Machine learning for climate change: Guiding discovery of sorbent materials for direct air capture of CO2. In *ICML 2021 Workshop on Tackling Climate Change with Machine Learning*, 2021. URL https://www.climatechange.ai/papers/icml2021/84.

Papenmeier, L., Nardi, L., and Poloczek, M. Increasing the scope as you learn: Adaptive Bayesian optimization in nested subspaces. In *Advances in Neural Information Processing Systems*, 2022.

Park, J. Contextual Bayesian optimization with trust region (CBOTR) and its application to cooperative wind farm control in region 2. *Sustainable Energy Technologies and Assessments*, 38:100679, 2020.

Park, J. and Law, K. H. A Bayesian optimization approach for wind farm power maximization. In *Smart Sensor Phenomena, Technology, Networks, and Systems Integration 2015*, volume 9436, pp. 46–57. SPIE, 2015.

Park, J. and Law, K. H. Bayesian ascent: A data-driven optimization scheme for real-time control with application to wind farm power maximization. *IEEE Transactions on Control Systems Technology*, 24(5):1655–1668, 2016a.

Park, J. and Law, K. H. A data-driven, cooperative wind farm control to maximize the total power production. *Applied Energy*, 165:151–165, 2016b.

Park, J., Kwon, S.-D., and Law, K. H. A data-driven approach for cooperative wind farm control. In *2016 American Control Conference (ACC)*, pp. 525–530. IEEE, 2016.

Park, J., Kwon, S.-D., and Law, K. A data-driven, cooperative approach for wind farm control: a wind tunnel experimentation. *Energies*, 10(7):852, 2017.

Park, J., Ferguson, M., and Law, K. H. Data driven analytics (machine learning) for system characterization, diagnostics and control optimization. In *Workshop of the European Group for Intelligent Computing in Engineering*, pp. 16–36. Springer, 2018.

Pedersen, M. M., van der Laan, P., Friis-Møller, M., Rinker, J., and Réthoré, P.-E. Dtuwindenergy/pywake: Pywake. *Zenodo [code]*, 10, 2019.

Pradhan, B., Sameen, M. I., Al-Najjar, H. A., Sheng, D., Alamri, A. M., and Park, H.-J. A meta-learning approach of optimisation for spatial prediction of landslides. *Remote Sensing*, 13(22):4521, 2021.

Rainforth, T., Foster, A., Ivanova, D. R., and Smith, F. B. Modern Bayesian experimental design. *arXiv preprint arXiv:2302.14545*, 2023.

Ramesh, S. S., Sessa, P. G., Krause, A., and Bogunovic, I. Movement penalized Bayesian optimization with application to wind energy systems. *arXiv preprint arXiv:2210.08087*, 2022.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., et al. Tackling climate change with machine learning. *ACM Computing Surveys (CSUR)*, 55(2):1–96, 2022.

Samaniego, F. P., Reina, D. G., Marín, S. L. T., Arzamendia, M., and Gregor, D. O. A Bayesian optimization approach for water resources monitoring through an autonomous surface vehicle: The Ypacarai lake case study. *IEEE Access*, 9:9163–9179, 2021.

Sanchez-Pi, N., Marti, L., Abreu, A., Bernard, O., de Vargas, C., Eveillard, D., Maass, A., Marquet, P. A., Sainte-Marie, J., Salomon, J., et al. Artificial intelligence, machine learning and modeling for understanding the oceans and climate change. In *NeurIPS 2020 Workshop-Tackling Climate Change with Machine Learning*, 2020.

Severson, K. A., Pfeiffer, O. P., Chen, J., Gong, K., Gregory, J., Goodwin, R., and Olivetti, E. A. Amortized inference of Gaussian process hyperparameters for improved concrete strength trajectory prediction. *Tackling Climate Change with Machine Learning: workshop at NeurIPS 2021*, 2021.

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and de Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE*, 104(1):148–175, 2015. ISSN 1558-2256. doi: 10.1109/JPROC.2015.2494218.

Singh, A., Ramos, F., Whyte, H. D., and Kaiser, W. J. Modeling and decision making in spatio-temporal processes for environmental surveillance. In *2010 IEEE International Conference on Robotics and Automation*, pp. 5490–5497, May 2010. doi: 10.1109/ROBOT.2010.5509934.

Song, L., Xue, K., Huang, X., and Qian, C. Monte carlo tree search based variable selection for high dimensional Bayesian optimization. *arXiv preprint arXiv:2210.01628*, 2022.

Spooner, T., Jones, A. E., Fearnley, J., Savani, R., Turner, J., and Baylis, M. Bayesian optimisation of restriction zones for bluetongue control. *Scientific reports*, 10(1): 1–18, 2020.

Sun, S., Tiihonen, A., Oviedo, F., Liu, Z., Thapa, J., Zhao, Y., Hartono, N. T. P., Goyal, A., Heumueller, T., Batali, C., et al. A data fusion approach to optimize compositional stability of halide perovskites. *Matter*, 4(4):1305–1322, 2021.

Tabakhi, A. M., Yeoh, W., and Fioretto, F. The smart appliance scheduling problem: A Bayesian optimization approach. In *International Conference on Principles and Practice of Multi-Agent Systems*, pp. 100–115. Springer, 2020.

Thebelt, A., Tsay, C., Lee, R., Sudermann-Merx, N., Walz, D., Shafei, B., and Misener, R. Tree ensemble kernels for Bayesian optimization with known constraints over mixed-feature spaces. *Advances in Neural Information Processing Systems*, 35:37401–37415, 2022.

Tillmann, P., Jäger, K., and Becker, C. Minimising the levelised cost of electricity for bifacial solar panel arrays using Bayesian optimisation. *Sustainable Energy & Fuels*, 4(1):254–264, 2020.

Tran, K. and Ulissi, Z. W. Active learning across intermetallics to guide discovery of electrocatalysts for CO2 reduction and H2 evolution. *Nature Catalysis*, 1(9):696–703, 2018.

Trivedi, R. and Khadem, S. Peak demand management and schedule optimisation for energy storage through the machine learning approaches. In *IEEE EUROCON 2021-19th International Conference on Smart Technologies*, pp. 157–163. IEEE, 2021.

Tu, B., Gandy, A., Kantas, N., and Shafei, B. Joint entropy search for multi-objective Bayesian optimization. *arXiv preprint arXiv:2210.02905*, 2022.

Valentin, S., Kleinegesse, S., Bramley, N. R., Seriès, P., Gutmann, M. U., and Lucas, C. G. Designing optimal behavioral experiments using machine learning, 2023.

Watson-Parris, D. Machine learning for weather and climate are worlds apart. *Philosophical Transactions of the Royal Society A*, 379(2194):20200098, 2021.

Williams, C. K. and Rasmussen, C. E. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.

Wu, X., Gomes-Selman, J., Shi, Q., Xue, Y., Garcia-Villacorta, R., Anderson, E., Sethi, S., Steinschneider, S., Flecker, A., and Gomes, C. Efficiently approximating the pareto frontier: hydropower dam placement in the amazon basin. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Yang, S., Deng, X., Ti, Z., Yan, B., and Yang, Q. Cooperative yaw control of wind farm using a double-layer machine learning framework. *Renewable Energy*, 193: 519–537, 2022.

Zhan, S., Wichern, G., Laughman, C., Chong, A., and Chakrabarty, A. Calibrating building simulation models using multi-source datasets and meta-learned Bayesian optimization. *Energy and Buildings*, pp. 112278, 2022.

Zhang, X., Lu, D., Dong, H., Zhao, X., Brennan, F., and Liang, Y. Vibration suppression of multi-component floating structures via passive TMDs and Bayesian ascent. *Ocean Engineering*, 259:112088, 2022.

Zhong, M., Tran, K., Min, Y., Wang, C., Wang, Z., Dinh, C.-T., De Luna, P., Yu, Z., Rasouli, A. S., Brodersen, P., et al. Accelerated discovery of CO2 electrocatalysts using active machine learning. *Nature*, 581(7807):178–183, 2020.