Future Energy Systems Center Pre-publication Paper

This is a pre-publication paper under review by a journal. It can only be shared within Future Energy Systems Center sponsor companies and cannot be made public. When sharing within your sponsor company, please state that this paper cannot be shared outside of the company.

MITei
MIT Energy Initiative

Learning-assisted Stochastic Capacity Expansion Planning: A Bayesian Optimization Approach

Aron Brenner, Rahman Khorramfar, Dharik Mallapragada, Saurabh Amin

arXiv:2401.10451v3 [eess.SY] 24 Jan 2024

***Abstract*—Solving large-scale capacity expansion problems (CEPs) is central to cost-effective decarbonization of regional-scale energy systems. To ensure the intended outcomes of CEPs, modeling uncertainty due to weather-dependent variable renewable energy (VRE) supply and energy demand becomes crucially important. However, the resulting stochastic optimization models are often less computationally tractable than their deterministic counterparts. Here, we propose a learning-assisted approximate solution method to tractably solve two-stage stochastic CEPs. Our method identifies low-cost planning decisions by constructing and solving a sequence of tractable temporally aggregated surrogate problems. We adopt a Bayesian optimization approach to searching the space of time series aggregation hyperparameters and compute approximate solutions that minimize costs on a validation set of supply-demand projections. Importantly, we evaluate solved planning outcomes on a held-out set of test projections. We apply our approach to generation and transmission expansion planning for a joint power-gas system spanning New England. We show that our approach yields an estimated cost savings of up to 3.8% in comparison to benchmark time series aggregation approaches.**

***Index Terms*—planning, power systems, stochastic optimization, mixed integer linear programming, machine learning, uncertainty**

I. Introduction

A. Motivation

The transition from current fossil fuel-dominated energy systems to deeply decarbonized ones requires coordinated infrastructure planning and operations while accounting for uncertainties in key operational parameters that capture availability of weather-dependent variable renewable energy (VRE) supply and energy demand. Increasing penetration of VREs – and increasing electrification of end-uses – contribute to these uncertainties, with significant implications for energy infrastructure planning [1], [2]. Additionally, interdependencies between electric power and natural gas (NG) systems, as primary energy vectors, are intensified owing to the shifting role of NG-based generation to compensate for the intermittent nature of VREs, substitution of gas with electricity in end-uses (e.g. heating), and emerging technologies such as NG-based generation with carbon capture and storage [3].

Capacity expansion problems (CEPs) form a crucial part of the energy systems planning toolkit as they guide infrastructure investment needed to meet future demand and decarbonization targets. Here, we focus on CEPs for low-carbon energy systems that incorporate multiple features. First, we account for the stochastic and weather-sensitive nature of key operational parameters in CEP models such as energy demand and VRE supply potential. Second, we recognize the importance of embedding operational dynamics such as ramping and storage at an hourly temporal fidelity. This becomes crucial as primary energy supply shifts to VRE generation and demand becomes more flexible [4], [5]. Finally, we consider the underlying network structure of energy vectors and their associated operating constraints in our CEP model to avoid underestimation of system cost [6].

While including these aforementioned features allows us to more realistically capture system operations in CEP formulations, the resulting model is rendered computationally intractable due to its large-scale and mixed-integer nature. Even for simpler CEPs that omit these features, the literature on energy systems planning makes simplifying assumptions to trade-off model fidelity for computational tractability [7], [8]. Most commonly, planners resort to *time series aggregation*, in which the CEP model is solved for a *representative operational period* that (hopefully) spans a wide range of supply and demand patterns observed over the full operating horizon. Solving the resulting reduced-order CEP alleviates computational burden and can yield solutions that are comparable or similar to the solution of the CEP over the full planning horizon [9], [10]. Unfortunately, these reduced-order CEPs do not account for the stochastic nature of input data – particularly VRE resource potential and energy demand – and thus have limited value from a planning perspective where cost effectiveness and constraint satisfaction under future scenarios are important criteria. On the other hand, solving full-scale stochastic CEPs with multiple scenarios of VRE and demand realization becomes computationally intractable with only a handful of scenarios, particularly when considering discrete investment decisions. This challenge necessitates a new approach that integrates the stochastic nature of VRE and demand parameters in CEPs while remaining computationally scalable.

B. Time Series Aggregation

Time series aggregation for energy systems has a rich literature [10], [11]. In the context of planning

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

problems, researchers select representative periods based on electric power demand and/or capacity factors (CFs) for VREs [12], [13], [14], [9]. To automatically select representative periods from this supply-demand data, one solves a *representative period clustering* (RPC) problem using well-known clustering algorithms such as k-means [15], [12], [13], k-medoids [12], [13], [9], or hierarchical clustering [16], [14].

Although clustering approaches to time series aggregation partially address the computational challenge in solving CEPs at scale, they face two major shortcomings. First, hyperparameters of the RPC problem are often selected in an ad hoc manner; these include the number of representative periods and the relative weights of various supply and demand features in the distance metric used for clustering. This precludes exploring a range of potential CEP instances that can differ in terms of inputs such as network topology, resource availability, and demand patterns. Additionally, hyperparameters selected for one CEP instance may not yield effective planning outcomes for other instances. Second, despite the well-known fact that weather variation impacts both supply and demand projections [15], typical approaches to time series aggregation largely ignore this uncertainty in selecting representative periods. Instead, the performance of clustering methods is simply evaluated by quantifying their ability to reproduce the CEP investment decisions using a *single* supply-demand projection [13], [17], [18]. This yields planning outcomes that are tailored to a specific projection of supply-demand data but might be hugely suboptimal for supply-demand patterns that might be realized for weather-years in the planning horizon. In effect, existing clustering methods applied to CEPs may produce investment outcomes that are not optimal under inter-annual variability. As weather-dependent VRE supply is expected to grow with decarbonization efforts, there is growing interest from researchers [19], [20] and system operators [21] to identify CEP investment outcomes that *meet demands at low cost under uncertain VRE supply and energy load profiles*.

Fig. 1 highlights the implications of these shortcomings in the case of a CEP for joint power-NG system planning [3]. We observe that investment decisions that minimize costs for a single supply-demand projection can yield much higher operational costs for different realization of demands and VRE availabilities (as seen by out-of-sample projections shown). We also notice that average costs incurred over 14 *out-of-sample* projections (i.e., supply-demand projections that are not used to instantiate the model) do not decrease monotonically with the number of representative days and are instead minimized at 25 representative days. When using more than 25 representative days, high out-of-sample costs may result from "overfitting" of investment decisions to the single projection used to instantiate the *surrogate* (i.e., reduced-order) CEP. This phenomenon necessitates careful tuning of RPC hyperparameters with consideration of *multiple* supply-demand projections and systematic *out-of-sample cost evaluations*. Our work addresses this challenge using a learning-assisted approach to time series aggregation.

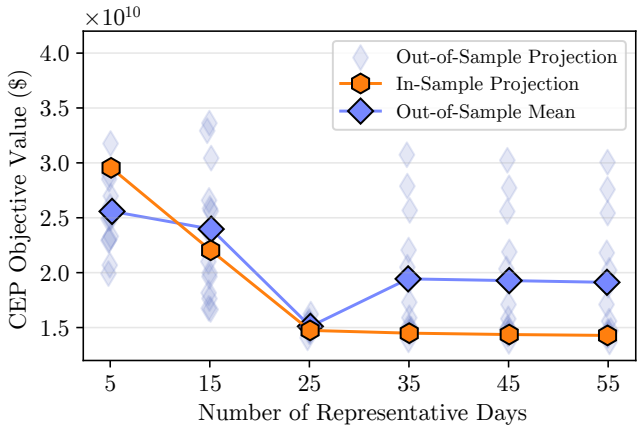


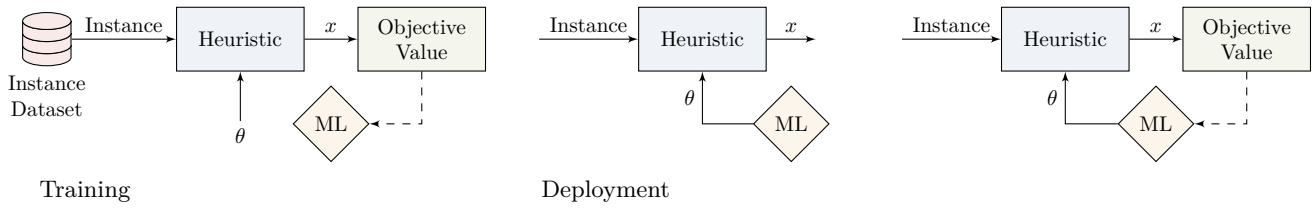
Fig. 1: Effect of varying the number of representative days on the CEP objective evaluated over 14 out-of-sample supply-demand projections using a single projection solution for a joint power-gas CEP (Sec. IV).

C. Contributions

We consider two-stage stochastic CEP for coordinated planning of multiple energy vectors. We develop a learning-assisted approach to compute investment decisions that yield the lowest combined investment and operational cost in expectation over multiple year-long projections of energy demand and VRE availability (Sec. II). Rather than solve an extensive form stochastic program, which is intractable due to the large-scale and combinatorial nature of the problem, our approach designs and solves a surrogate CEP instantiated using a smaller set of representative periods. Note, however, that there can be many potential surrogate CEPs, each yielding different planning outcomes. Our key contribution is to leverage a Bayesian optimization (BO) approach to efficiently search the space of RPC hyperparameters. We demonstrate that these automatically tuned hyperparameters help identify planning outcomes with the lowest expected total cost (investment plus operational) over *out-of-sample* supply-demand projections (Sec. III).

Our approach exploits two properties shared by many CEPs for energy systems planning: (1) while optimizing investment decisions for a full planning horizon is difficult (i.e., mixed-integer), evaluating the operational cost incurred by a fixed set of investment decisions often reduces to solving a linear program for each projection; (2) investment decisions obtained from solving a surrogate problem instantiated over a

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.



(a) Learning-assisted heuristics (b) Our BO-assisted approach

Fig. 2: Conceptual difference between offline learning-assisted heuristic approaches in training and deployment (a) and our proposed BO-assisted approach (b). Here, θ denotes the heuristic hyperparameters, x denotes the decision variables of the optimization task, and ML denotes the machine learning module (i.e., BO in our approach). Dashed lines indicate the flow of decision costs to ML model estimation.

small number of representative periods can yield more robust planning outcomes (in terms of better out-of-sample performance) with less computational effort in comparison to solving a larger surrogate problem. This second property has not been explored in the existing literature. We show how these properties can be systematically leveraged to tractably solve two-stage stochastic CEPs.

Our BO-assisted approach can be viewed as a "learning to configure" approach [22] in that it learns to select hyperparameters specifying the time series aggregation heuristic for surrogate model construction. In this regard, our approach is distinct from recent works that use offline learning to improve optimization heuristics [23], [24] as it can be deployed *without pre-training* on a large number of problem instances (Fig. 2). Specifically, our approach quickly identifies low-cost investment decisions by learning to identify promising hyperparameter configurations for a *single problem instance* through an iterative process of instantiating surrogate problems, solving them, and evaluating the resulting investment decisions over a range of validation supply-demand projections. We demonstrate the effectiveness of this approach for joint power-gas generation and transmission expansion planning for New England with 20 yearly supply-demand projections (Sec. IV).

II. Problem Formulation

A. Capacity Expansion Problem

We consider a stochastic generalization of two-stage CEPs that commonly arise in energy systems planning:

$$[\text{CEP}] \quad \min_{x} \quad c^\top x + \mathbb{E}_{\omega\in\Omega}[Q(x,\omega)] \tag{1a}$$

$$\text{s.t.} \quad Ax = b. \tag{1b}$$

The first-stage variable x takes integer and continuous values and denotes the investment or decommissioning decisions that are realized before the planning horizon. These describe locations and investment levels for various plant types (e.g., solar panels, wind turbines, gas-fired plants), transportation (i.e., transmission lines for power systems, pipelines for NG systems), and storage facilities. Constraints (1b) imposes first-stage constraints such as budget and number of operational assets. The objective function (1a) minimizes the combined investment and operational cost over a scenario set, Ω.

For a feasible first-stage solution x, the value function of the recourse problem for each year-long stochastic scenario (i.e., supply-demand projection) $\omega \in \Omega$ is given by

$$Q(x,\omega) = \min_{y} \quad \sum_{t\in\mathcal{T}} d_t^\top y_t \tag{2a}$$

$$\text{s.t.} \quad B_t^\omega x + C_t y_t = q_t^\omega \qquad t \in \mathcal{T} \tag{2b}$$

$$D x + \sum_{t\in\mathcal{T}} E y_t = p, \tag{2c}$$

where y_t is a continuous variable that denotes second-stage decisions in period t. Without loss of generality, we assume that each $t \in \mathcal{T}$ corresponds to a single day of the planning period. The objective function (2a) minimizes operational costs, which can include variable generation, load shedding penalty and fuel costs. The constraints (2b) link capacity expansion decisions in the first stage to operational decisions in the second stage. In particular, these constraints ensure energy balance of demands q_t^ω, load shedding y_t, and generation constrained by first-stage decision variables and renewable capacity factors B_t^ω. The second set of constraints (2c) links operational decisions across periods such as ramping and storage for adjacent hours and cross-sectoral emissions limits throughout the entire planning horizon. Here, all inequality constraints are formulated as equality constraints using slack variables. Fig. 3 illustrates how supply and demand projections enter into the problem to form the constraint matrix.

We assume that (2) is feasible for any first-stage decisions, x, i.e., load shedding is unconstrained and incurred as a cost. Additionally, we note that each period, $t \in \mathcal{T}$, does not necessarily correspond to the planning resolution of energy systems. For instance, while hourly time resolution is usually used for electric power planning, daily operating resolution can be more suitable

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

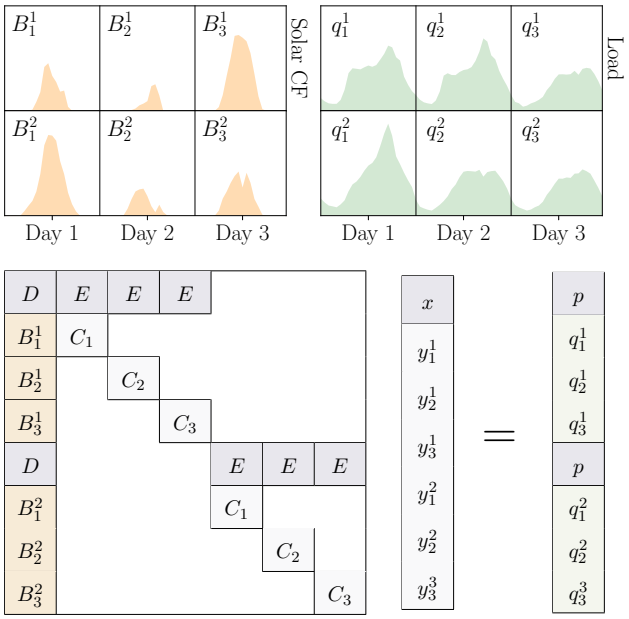


Fig. 3: Supply-demand projections [25] (top) and constraints (bottom) for an illustrative stochastic capacity expansion problem with two projections and three periods (i.e., days).

for gas system planning [3]. Following [13], [26], [27], we assume representative periods to denote daily time spans.

B. Surrogate Capacity Expansion Problem

In practice, the number of second-stage decision variables and constraints, $\mathcal{O}(|\Omega||\mathcal{T}|)$, is orders of magnitude larger than those corresponding to the first stage. This renders an intractable formulation even for CEPs that consider a single scenario (on the order of 10^6 variables and constraints in our computational study with an hourly resolution). As a result, scenario-wise decomposition methods cannot be applied to solve (1) quickly. Furthermore, applying a day-by-day decomposition approach requires relaxing inter-day linking constraints [28]. Consequently, we must consider approximate solution algorithms for (1). Such an approach instead solves a surrogate model – a reduced-order formulation of (1) instantiated on a representative day set, $\mathcal{T}_R \subset \mathcal{T}$, with a reduced number of variables and constraints, $\mathcal{O}(|\Omega||\mathcal{T}_R|)$.

To formulate our surrogate model, we first construct the unordered sample set $\mathcal{S}$, which contains daily supply-demand profiles for the entire planning horizon for all scenarios. Specifically, for a planning horizon of one year, $\mathcal{S}$ contains $365 \times |\Omega|$ samples. This sample set $\mathcal{S}$ is distinguished from the notion of a scenario set as a scenario $\omega \in \Omega$, describes a *year-long* supply-demand projection. On the other hand, the sample set describes the unordered collection of *day-long* supply-demand projections taken from across all scenarios.

Next, we select θ_R days from $\mathcal{S}$ and construct the set of representative days $\mathcal{T}_R$. Additionally, we assign a weight $w_t \in \mathcal{W}$ to each representative day in $\mathcal{T}_R$ such that $\sum_{t\in\mathcal{T}_R} w_t = |\mathcal{T}|$. The resulting surrogate problem can be formulated as the single-scenario CEP,

$$
\begin{aligned}
[\mathrm{CEP_S}(\mathcal{T}_R,\mathcal{W})] \quad \min_{x,y} \quad & c^\top x + \sum_{t\in\mathcal{T}_R} w_t d_t^\top y_t && \text{(3a)}\\
\text{s.t.} \quad & Ax = b && \text{(3b)}\\
& B_t^\omega x + C_t y_t = q_t^\omega, \quad t \in \mathcal{T}_R && \text{(3c)}\\
& Dx + \sum_{t\in\mathcal{T}_N} w_t E y_t = p. && \text{(3d)}
\end{aligned}
$$

We note that different orderings of days in $\mathcal{T}_R$ instantiate different surrogate problems. For simplicity, we assume that days in $\mathcal{T}_R$ are arranged in an increasing order of the scenario index and day of year.

C. Representative Period Clustering Problem

The choice of representative days is crucial to achieving a computationally light surrogate model that well-approximates (1). In other words, the investment decisions x resulting from $\mathrm{CEP_S}(\mathcal{T}_R,\mathcal{W})$ should approximately minimize (1a). Existing literature, including [29], [26], [14], [12], propose clustering-based approaches for identifying such representative periods. Here, we consider the following k-medoids clustering model, which we refer to as the representative period clustering (RPC) problem:

$$
\begin{aligned}
[\mathrm{RPC}(\theta)] \quad \min \quad & \sum_{i=1}^{k}\sum_{t\in\mathcal{C}_i} D(m_i,t) && \text{(4a)}\\
\text{s.t.} \quad & \bigcup_{i=1}^{k} \mathcal{C}_i = \mathcal{S} && \text{(4b)}\\
& \mathcal{C}_i \cap \mathcal{C}_j = \emptyset \qquad \forall i = [\theta_R], j \neq i && \text{(4c)}\\
& (m_1,\ldots,m_{\theta_E}) = \mathcal{T}_E && \text{(4d)}\\
& \mathcal{C}_i = \{t\} \qquad \forall t \in \mathcal{T}_E, && \text{(4e)}
\end{aligned}
$$

where $\mathcal{C} := (\mathcal{C}_1,\ldots,\mathcal{C}_k)$ denotes the representative day clusters with corresponding medoids $(m_1,\ldots,m_k)$, $D(m_i,t)$ denotes the "distance" between medoid m_i and period t, and $\mathcal{T}_E \subset \mathcal{T}_R$ denotes a set of "extreme days," or days exhibiting particularly hard-to-satisfy demands, with corresponding weights $w_i = 1$ for all $i \in \mathcal{T}_E$. One can apply an out-of-the-box clustering algorithm [30] to solve (4) and obtain the representative day set and weights respectively as $\mathcal{T}_R = (m_1,\ldots,m_{\theta_R})$ and $\mathcal{W} = (\frac{1}{|\mathcal{S}|}|\mathcal{C}_1|,\ldots,\frac{1}{|\mathcal{S}|}|\mathcal{C}_{\theta_R}|)$. Here, we focus on the central issue of selecting hyperparameters for the RPC problem.

Before proceeding, we note that (4d) ensures that the representative day set $\mathcal{T}_R$ contains θ_E "extreme" days, $\mathcal{T}_E$, or days with high energy demands. Additionally, constraint (4e) enforces that each extreme day forms a single-member cluster. Including extreme days in the surrogate model may compensate for the smoothing effect of clustering and result in a planning outcome that is

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

more robust to peak demands. However, selecting the hyperparameter θ_E is not obvious a priori and again requires tuning to avoid over- or under-investment.

The distance function, $D(m_i, t)$ is another choice that greatly impacts the RPC solution, and consequently, planning outcomes. A natural choice is the Euclidean distance between supply-demand parameters associated with a pair of days,

$$D(t,t') = \theta_B \|B_t^\omega - B_{t'}^\omega\|_F + (1-\theta_B)\|q_t^\omega - q_{t'}^\omega\|_2,$$

where $\|\cdot\|_F$ denotes the Frobenius norm. We consider $\theta_B \in [0,1]$ to be a hyperparameter that weighs the relative importance of demand (contained in q_t^ω) and supply-side features such as renewable capacity factors (contained in B_t^ω). Importantly, different choices of θ_B impact RPC outcomes by influencing the daily supply- or demand- patterns captured in the representative day set. Selecting θ_B is not obvious. More generally, more than two groups of parameters may need to be weighed when multiple VRE technologies and energy vectors are considered, which is the case for our computational study (Sec. IV).

Fig. 4 illustrates one challenge in choosing θ_B that results from correlations among supply-demand parameters. For example, clustering days according to, e.g., NG loads, will also capture a wide range of early morning/late evening solar availability due to highly negative correlations between NG demand and solar availability resulting from the time of year. However, wind availability patterns may not be proportionately represented in the surrogate problem due to minimal correlation with NG loads. In the presence of such correlations, an RPC problem that weighs all parameters equally may, in fact, fail to adequately capture variability of certain parameters and consequently yield suboptimal planning outcomes.

Essentially, RPC hyperparameters can singificantly impact planning outcomes obtained by solving CEP$_S$. Previous works do not directly deal with this issue and instead focus on evaluating outcomes of individual hyperparameter choices, namely θ_R [12], [15] and θ_E [13], for specific planning contexts/formulations. In particular, most works that cluster according to both supply- and demand-side parameters avoid applying any weighting and select θ_B implicitly [13], [15].

Here, we highlight another limitation of prior work, specifically the assumption that planning outcomes improve monotonically in θ_R [12], [14]. Consequently, their results suggest that the ideal choice of θ_R is a point at which planning objectives [12], [14] or clustering objectives [31] show diminishing returns with respect to computational runtime. In contrast, we find that this is not the case for out-of-sample evaluations, as illustrated in Fig. 1. In particular, the ideal choices of θ_R, θ_E, and θ_B depend on the problem and the available data that can be used to instantiate the surrogate problem and evaluate resulting costs out-of-sample. The difficulty of navigating this hyperparameter space, which varies widely across problem settings, necessitates developing a systematic approach to *jointly* tuning multiple hyperparameters in a cost-driven manner. Next, we describe and evaluate a BO-assisted approach for hyperparameter tuning based on resulting out-of-sample cost objectives.

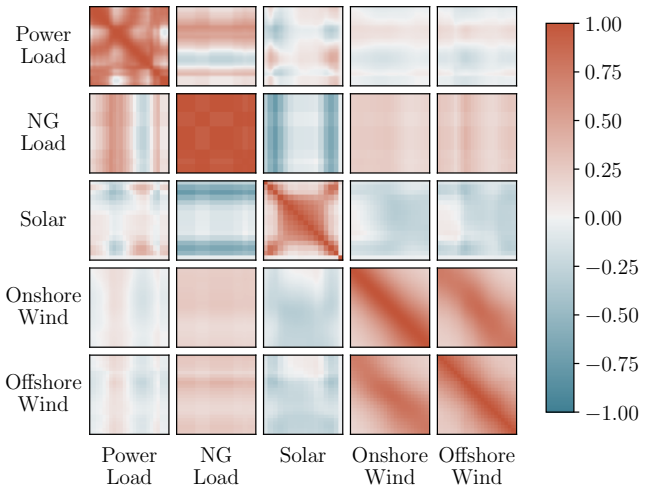


Fig. 4: Pearson correlations between supply-demand parameter groups for our computational study (Sec. IV). Each row/column of pixels corresponds to one hour of the day (averaged over system nodes) except for NG, where pixels correspond to daily nodal loads. For each pair of parameter groups, pixel (i,j) shows the correlation of parameter i in the first group with parameter j in the second group.

III. Solution Approach

To address this challenge systematically and for a range of planning contexts, we propose searching over a continuous space of RPC hyperparameters using Bayesian optimization (BO), a derivative-free search strategy for black-box functions that has been applied with success in hyperparameter tuning for ML applications [32]. At a high level, our BO-assisted approach learns to minimize $f(\cdot)$, the function mapping RPC hyperparameters to CEP objective values, by performing a series of *function evaluations*. As is the case in many BO applications, $f(\cdot)$ does not admit an analytic expression and can only be evaluated through a series of complex operations (i.e., solving optimization problems). In each iteration of our approach, a function estimate – a Gaussian process (GP) regression model – is re-estimated using the most recent function evaluation, and new candidate hyperparameters are identified according to their estimated potential for yielding a low CEP objective and improving the GP model estimate. Fig. 6 illustrates our solution approach and shows how hyperparameters θ are evaluated (Sec. III-A) and how subsequent candidate hyperparameters are selected (Sec. III-B) in each iteration.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

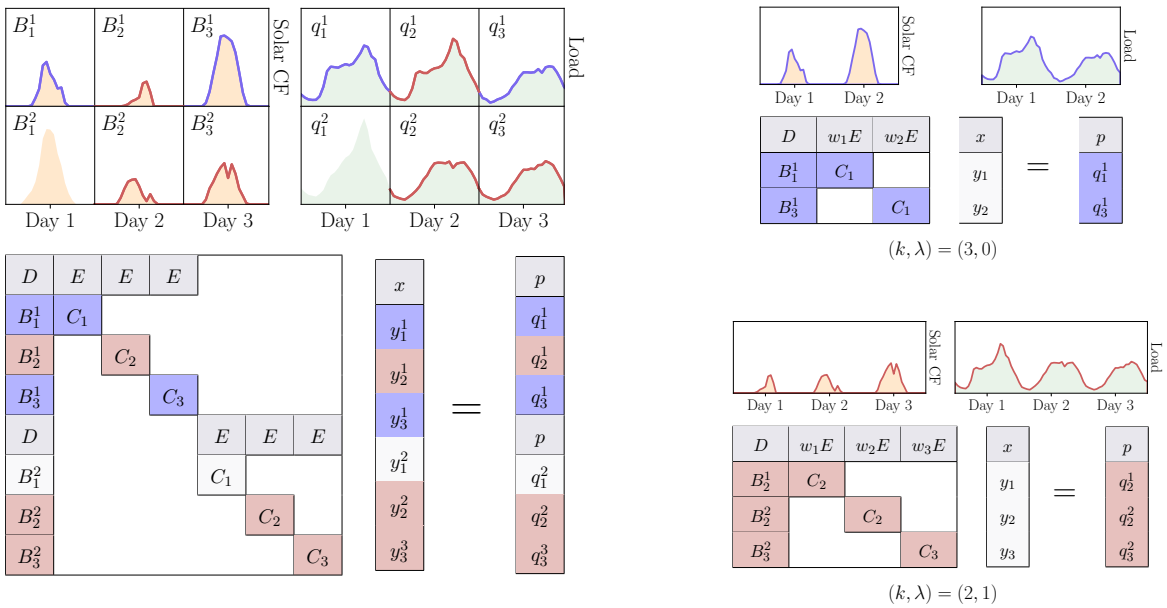


Fig. 5: Surrogate problems are instantiated as "single-scenario" CEPs with a representative day set constructed using days sourced from different supply-demand projections. Different choices of RPC hyperparameters instantiate different surrogate problems. Setting $(k, \lambda) = (2, 1)$, the resulting surrogate problem (top right) captures typical solar availability and nominal energy demand patterns. Setting $(k, \lambda) = (3, 0)$, the resulting surrogate problem (bottom right) captures a wider range of load profiles but fails to capture days with high solar availability.

A. Function Evaluation

Before describing the function evaluation process, we discuss the importance of dividing scenarios into disjoint sets of training and validation to reduce "look-ahead" bias. This is an optimistic bias in the estimate of operational cost that occurs as a result of including days from the second stage in the representative day set. In other words, if the scenario sets are not separated during training, the BO approach might learn to include specific days from the validation set that minimize validation costs. Consequently, the corresponding operational cost estimate will be lower, on average, than the actual operational cost incurred out-of-sample (assuming projections are drawn i.i.d.) as the representative day set has been "overfit" to the sample set (see Fig. 1). Accordingly, we split the scenario set Ω into training, validation, and test sets, $\Omega^{\text{train}}, \Omega^{\text{val}}, \Omega^{\text{test}}$ and evaluate costs over Ω^{val} in order to obtain a less biased estimate of costs incurred on Ω^{test} (i.e., out-of-sample).

Denoting the number of RPC hyperparameters by k, we define $\Theta \subset \mathbb{R}^k$ to be the bounded space of candidate RPC hyperparameters. We assume Θ to be continuous, which we generalize to integer-valued hyperparameters with rounding. In each iteration of the search, we evaluate $f(\theta)$ for some candidate hyperparameters $\theta \in \Theta$. For example, if we want to search over the space of $10 \leq \theta_R \leq 20$ and $1 \leq \theta_E \leq 10$, we define $\Theta := [10, 20] \times [1, 10]$ and in each iteration evaluate some candidate $\theta = (\theta_R, \theta_E)$. This function evaluation requires three steps.

1) **Clustering the sample set.** We solve (4) using a sample set constructed from Ω^{train} to identify a set of representative days (i.e., cluster medoids) and their corresponding weights, $\mathcal{T}_R$ and $\mathcal{W}$ respectively.

2) **Obtaining investments.** We then solve (3) for the given representative days to obtain investment decisions, x, and the corresponding investment cost, $c^\top x$.

3) **Evaluating operational costs.** Given the investment decisions x from the second step, we solve the resulting recourse problem (2) for the full planning horizon, $\mathcal{T}$, for all scenarios in Ω^{val}. This yields a set of operational decisions and corresponding operational costs for all validation scenarios. Averaging these values gives $f(\theta) = c^\top x + \frac{1}{|\Omega^{\text{val}}|} \sum_{\omega \in \Omega^{\text{val}}} \sum_{t \in \mathcal{T}} d_t^\top y_t^\omega$, an upper bound (i.e., feasible solution) for (1a) instantiated on Ω^{val}.

Regarding computational runtime, solutions for RPC(θ) can be obtained quickly with approximate algorithms such as PAM [30]. The computational burden in Step (3) is also likely to be low as it requires solving $|\Omega^{\text{val}}|$ recourse problems, which are linear programs in our case. Additionally, Step (3) can easily be parallelized to accelerate the process. In most cases, Step (2) of the function evaluation will require the greatest runtime as it involves solving a mixed-integer linear program (3) with $\mathcal{O}(|\mathcal{T}_R|)$ second-stage variables and constraints. Consequently, the worst-case runtime of one function evaluation depends on the maximum value that θ_R (i.e., the number of representative days) takes in Θ.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

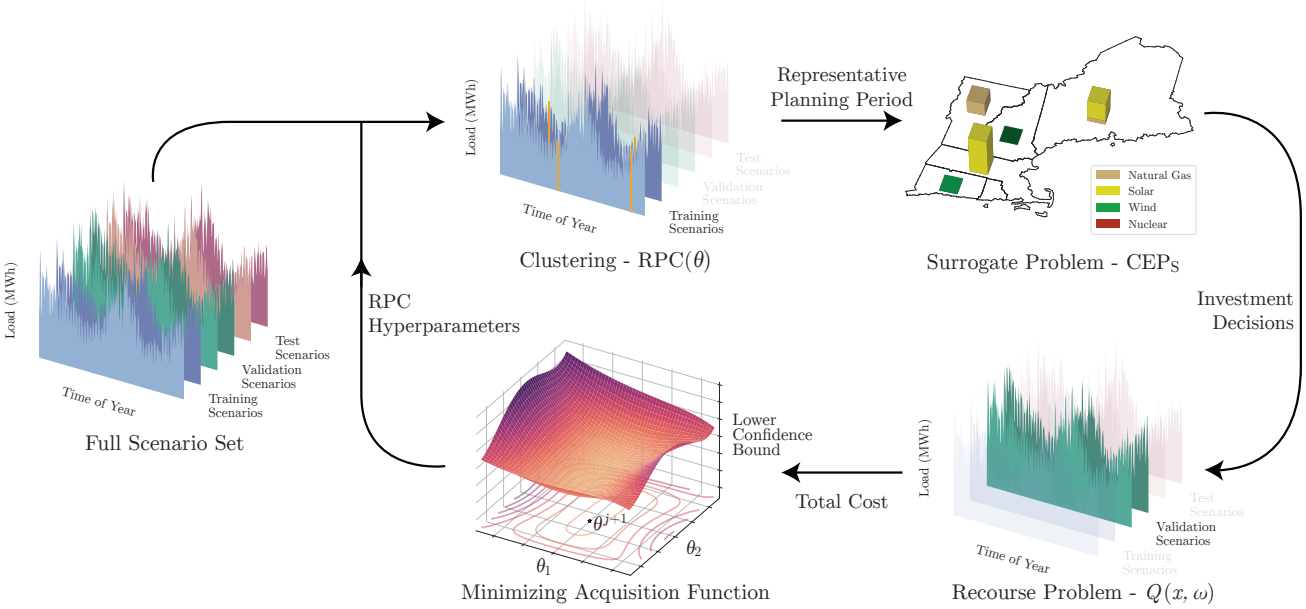


Fig. 6: The proposed BO-assisted solution approach.

B. Selecting Candidate Hyperparameters

Following each function evaluation, the procedure identifies the next candidate hyperparameters in three steps. Let us suppose that we have already evaluated j hyperparameter settings, $\theta^1, \ldots, \theta^j$, which has returned function values $f(\theta^1), \ldots, f(\theta^j)$.

1) **Instantiating a prior.** We instantiate a prior predictive covariance function, or *kernel function*, $\Sigma : \Theta \times \Theta \to \mathbb{R}$, which yields the GP prior distribution

$$f^{1:j} \sim \mathcal{N}(0, \Sigma(\theta^{1:j}, \theta^{1:j})), \tag{5}$$

where $\theta^{1:j}$ denotes $\theta^1, \ldots, \theta^j$ and the kernel function is defined as:

$$\Sigma(\theta, \theta') = \frac{2^{-3/2}}{\Gamma(5/2)}(\sqrt{5}\|\theta - \theta'\|_2)^{5/2} K_{5/2}(\sqrt{5}\|\theta - \theta'\|_2),$$

where $\Gamma(\cdot)$ and $K_{5/2}(\cdot)$ are the gamma function and modified Bessel function respectively [33]. This is a common choice of kernel function [33], [32] and is the default setting in BoTorch [34], the package that we used in our experiments. Note that in (5) we assume $f^{1:j}$ to have zero mean, which we enforce in practice through standardization.

2) **Performing a Bayesian update**. We then estimate the GP posterior conditional distribution of CEP validation costs over the hyperparameter space by applying a Bayesian update. This posterior is given by

$$f \mid \theta, \theta^{1:j}, f^{1:j} \sim \mathcal{N}(\mu(\theta), \sigma^2(\theta)),$$

where

$$\mu(\theta) = \Sigma(\theta, \theta^{1:j})\Sigma(\theta^{1:j}, \theta^{1:j})^{-1} f^{1:j} \tag{6}$$
$$\sigma^2(\theta) = \Sigma(\theta, \theta) - \Sigma(\theta, \theta^{1:j})\Sigma(\theta^{1:j}, \theta^{1:j})^{-1}\Sigma(\theta^{1:j}, \theta) \tag{7}$$

are the posterior predictive mean and variance functions respectively [35].

3) **Minimizing the acquisition function.** To identify promising candidate hyperparameters, we must define an *acquisition function* that quantifies potential for evaluation as a function of the estimated GP posterior. Accordingly, we consider the *lower confidence bound* (LCB) acquisition function

$$\text{[LCB]} \qquad a_{\text{LCB}}(\theta) = \mu(\theta) - \beta\sigma^2(\theta), \tag{8}$$

and in each iteration select θ^{j+1} that minimizes $a_{\text{LCB}}(\theta)$. Here, β is a GP hyperparameter (not to be confused with hyperparameters of the RPC problem) that tunes the exploration-exploitation tradeoff of selecting candidates with low predicted means as opposed to those with high predicted variances [32]. We note that the second stage will always be feasible for any choice of investment decisions, x, and by extension, RPC hyperparameters, θ. Consequently, (8) is unconstrained and can be solved quickly using quasi-Newton methods such as BFGS [36], which are implemented as part of standard Bayesian optimization packages.

Most implementations of BO approaches recommend initializing the GP model by evaluating a set of N_0 points, $\theta^1, \ldots, \theta^{N_0}$ sampled from some distribution over the search space Θ [33], [32]. Initializing the GP with a larger random sample improves the estimate of the GP but increases the number of function evaluations required before "higher potential" candidates can be evaluated. Properly normalizing $\theta^1, \ldots, \theta^j$ and $f(\theta^1), \ldots, f(\theta^j)$ before estimating the GP also has a significant impact on the performance of BO in practice. Following BoTorch specifications [34], we rescale Θ to $[0, 1]^m$ before estimation. Additionally, we standardize

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

the function evaluations so that $f(\theta^1), \ldots, f(\theta^j)$ has zero mean and unit variance in each iteration before estimating the GP.

IV. Computational Study

In this section, we evaluate the ability of our proposed BO-assisted approach to identify low-cost planning outcomes by searching the low-dimensional space of RPC hyperparameters for a stochastic variant of the Joint Power and Natural Gas (JPoNG) model [3]. The JPoNG model evaluates the cost-optimal planning of power and NG infrastructure with a resolved representation of spatial, temporal, and technological system constraints. The model structure follows that of a two-stage stochastic CEP that minimizes annualized investment and operating cost of bulk infrastructure of both vectors over representative days. The power system's operational constraints include ramping, unit commitment, state of charge for storage, and direct current (DC) power flow. Further details of the model are available in [3], [25].

Our computational experiments are based on a 6-node power system and 23-node NG system representation of the New England region, considering projected 2050 energy demand, technology cost assumptions, and an 80% decarbonization constraint with respect to 1990 levels [37]. Operational decisions and parameters in the power system exist at an hourly resolution while those in the NG system are at a daily resolution. The operational scenario set consists of 5 training, 10 validation, and 5 test projections for 2050. These stochastic scenarios contain parameters encoding both power-gas load and VRE capacity factors based on observations from 20 weather-years after assuming moderate electrification of residential heating [25].

We classify these parameters into five groups: (1) demand for electric power, (2) demand for NG, (3) solar CF, (4) onshore wind CF, and (5) offshore wind CF. Our BO approach tunes the relative weights of these parameter groups in the RPC distance function. Accordingly, we introduce hyperparameters $\theta_1, \ldots, \theta_5$ and bound $\theta_i \in [0, 1]$ for all i. Since relative distances are only unique up to re-scaling of $\theta_1, \ldots, \theta_5$, we impose that $\theta_1 + \cdots + \theta_5 = 1$. We also introduce two hyperparameters specifying the number of non-extreme days, $5 \leq \theta_6 \leq 80$, and the number of extreme days, $0 \leq \theta_7 \leq 10$. Specifically, we introduce θ_7 extreme days for both the power and NG systems according to total daily load so that $2 \times \theta_7$ extreme days are included in total.

Using BoTorch [34], we implement two BO approachs with different GP hyperparameters, $\beta = 10$ and $\beta = 2$ (Eq. 8). We denote these two methods by BO_{10} and BO_2 respectively. We conduct four trials, each of which corresponds to an initialization with $N_0 = 20$ random function evaluations followed by 80 iterations of our BO-assisted approach. The resulting hyperparameters are evaluated according to their corresponding validation cost, f^{val} (Sec. III-A), and test cost, f^{test}, defined to be the average cost incurred over the set of test projections. We also compare our approach to a random search heuristic that evaluates hyperparameters θ sampled uniformly at random from Θ.

For all trials, we also report percentage improvement over two baseline solutions selected to represent the common practice in the literature of selecting a high number of representative days and clustering with all parameters weighted equally [12], [14]. In both baselines, we fix $\theta_1 = \cdots = \theta_5 = 0.2$ and set $\theta_6 = 80$. The two baselines are distinguished according to the number of extreme days they consider: the first (Base. 1) sets $\theta_7 = 0$ while the second (Base. 2) sets $\theta_7 = 10$. In our experiments, we also obtained results for a third baseline that uses $\theta_7 = 2$ extreme days and obtains the highest average validation and test costs among all results ($14.38 and $14.59 billion respectively). However, we omit this baseline from discussion of results due to space constraints.

A. Cost Comparison

Fig. 7a shows convergence of all three methods averaged over the four trials. The random search heuristic shows relatively slow convergence as it is not able to incorporate previous function evaluations to identify promising regions of the search space with effective capacity expansion decisions (or conversely, to avoid regions with ineffective capacity expansion decisions). On the other hand, BO_{10} quickly converges to low-cost solutions. BO_2 also outperforms random search but converges more slowly and to higher-cost solutions. This might reflect the non-smooth nature of f in small neighborhoods due to planning outcomes being sensitive to small changes to θ. Consequently, an acquisition function that favors exploration (i.e., the one used for BO_{10}) outperforms one that evaluates multiple θ in a small neighborhood after obtaining one low-cost solution.

Table I shows the best hyperparameter configuration identified for each method over all trials and its corresponding cost and improvement over the baselines. Interestingly, test costs are lower than validation costs across most experiments. This is a reflection of the relative harder-to-meet supply-demand projections that exist in the validation set rather than a statistical bias intrinsic to the optimization procedure. Importantly, all search methods improve over the baseline test costs by at least 3% despite using fewer representative days. This demonstrates the necessity of tuning RPC hyperparameters over instantiating a large surrogate problem using generic (i.e., untuned) distance metrics.

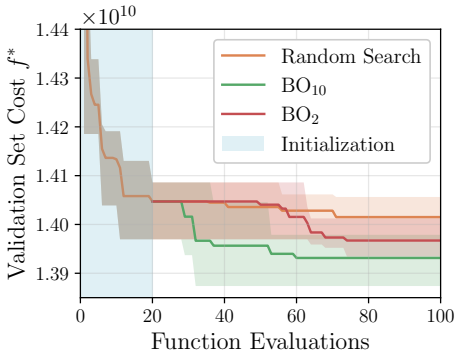
B. Hyperparameter Values

The best hyperparameter configuration – as evaluated by both validation and test costs – is identified by BO_{10}. This configuration assigns relatively high weight to variability in solar availability and comparatively

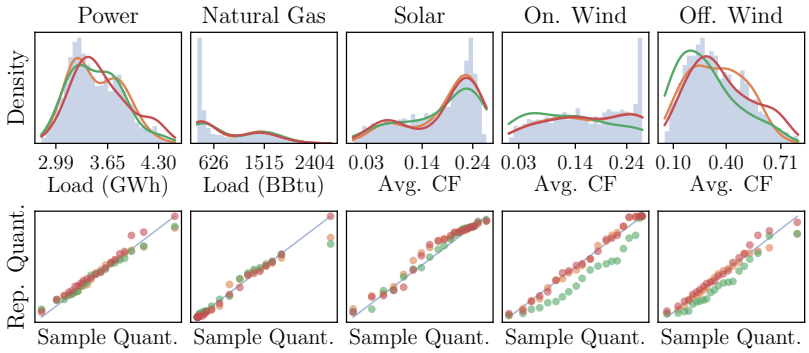
Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

TABLE I: Hyperparameters, objective values, and improvement over baseline costs for each method (corresponding to lowest test cost across four trials). Percentage improvement is calculated with respect to the best performing baseline solution for the respective scenario set (i.e., Base. 1 for the validation set and Base. 2 for the test set). Dashes denote zero values.

	Aggregation Hyperparameters (θ^*)							Cost ($ billion)		Improv. (%)	
Method	Power	NG	Solar	Off.	On.	Rep.	Ext.	f^{val}	f^{test}	Val.	Test
Base. 1	0.2	0.2	0.2	0.2	0.2	80	–	14.35	14.41	–	-0.29
Base. 2	0.2	0.2	0.2	0.2	0.2	80	10	14.42	14.37	-0.5	–
Random Search	0.367	0.106	0.103	0.333	0.091	57	–	13.97	13.93	2.7	3.0
BO ($\beta = 10$)	0.389	–	0.611	–	–	46	–	**13.87**	**13.82**	3.3	3.8
BO ($\beta = 2$)	–	0.318	0.269	0.414	–	64	2	13.93	13.89	2.9	3.4



(a) Lines show lowest validation costs obtained thus far averaged over all trials while shaded regions show highest and lowest costs obtained over all trials. 'Initialization' denotes the random search phase used to estimate initial GP models.



(b) Top: histograms of daily parameter values observed in the sample set (blue) and in the rep. day sets corresponding to Table I (shown as kernel density estimates). Values are averaged over hours and nodes for all parameters except gas load. Bottom: quantile-quantile plots comparing the empirical distribution of the sample set to the identified rep. day sets.

Fig. 7: Clustering and objective cost results corresponding to the lowest-cost trial for each search method.

lower weight to offshore, and particularly onshore, wind availability. As a result, the distribution of daily power loads and solar CFs in the representative day set closely approximates the distribution observed in the larger sample set (Fig. 7b). This is not the case for wind availability, which is largely underestimated by the representative day set. On the other hand, random search and BO_2 both learn to weight offshore wind CFs highly, and as a result, closely approximate the distribution of daily offshore wind availabilities observed in the sample set. Additionally, these methods also well-approximate the distribution of daily *onshore* wind availabilities, which is reflective of the high correlation between onshore and offshore wind availabilities seen in Fig. 4.

While we observe that generally at least one of either power or NG load is weighted highly, the total weight assigned to these parameters is less than that assigned to VRE CFs. This reflects the relative variability of VRE availability profiles, which are highly intermittent hour-to-hour, as compared to energy demand profiles. Moreover, demand for power is relatively correlated with demand for NG (Fig. 4). As a result, a sufficient range of power and NG demand patterns will likely be captured as long as at least one of the two parameters is weighted highly during clustering (Fig. 7b). This explains the weight of 0 assigned to NG by BO_{10}, which nevertheless obtains low-cost planning outcomes.

The identified hyperparameter configurations also include at least 46 representative days but no more than 2 extreme days (if any). The absence of explicitly chosen extreme days in many of the instances indicates that the clustering approach sufficiently captures load volatility in the manner it partitions the planning horizon. In line with Fig. 1, in which solutions were obtained from clustering a single scenario (with $\theta_1 = \cdots = \theta_5 = 0.2$), we observe that the selected number of representative days is less than the maximum number allowed. This again demonstrates that, in settings considering out-of-sample operational cost objectives, increasing the number of representative days does not necessarily lead to lower

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

costs. On the other hand, the number of representative days chosen is greater than 25, which yielded the lowest average out-of-sample cost for the case shown in Fig. 1. This suggests that tuning distance weight hyperparameters and utilizing a larger sample set might help to reduce overfitting of planning outcomes to $\mathcal{T}_R$ when the number of representative days is high.

C. New Assets and Utilization

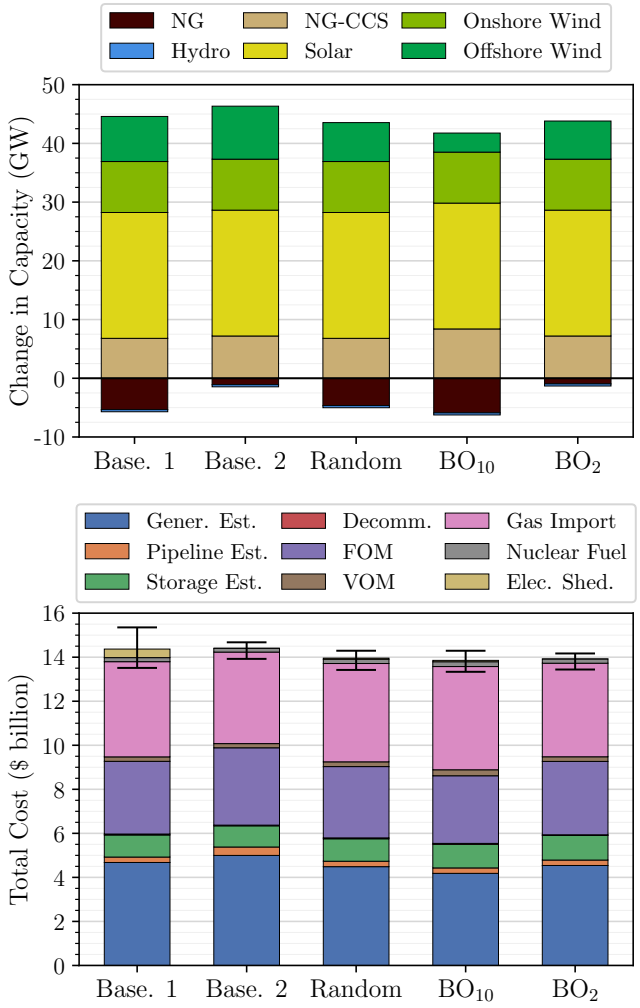


Fig. 8: Capacity change for power plants (top) and cost breakdown (bottom). Error bars show maximum and minimum total cost incurred over the 15 projections that include validation and test sets. 'NG' denotes existing gas-fired plants. 'NG-CCS' is a gas-fired plant with carbon capture and storage technology. 'FOM' and 'VOM' are fixed and variable costs for plants.

Fig. 8 shows installed capacity for power plants and cost components for all solved investment strategies. Comparing the two baselines, we find that including extreme days reduces costs from load shedding as well as inter-annual variations in operational costs (as shown by the narrower error bars for Base. 2). This is achieved by investing more aggressively in offshore wind generation while retaining a larger number of existing NG plants. Ultimately, this yields slightly higher costs on average, but significantly lower costs for worst-case projections as shown by the wider error bars for Base. 1 in Fig. 8.

Of all methods, BO_{10} yields the lowest average system cost. The resulting planning outcomes recommend high decommissioning and relatively low capacity addition for wind power (both onshore and offshore) compared to all other approaches. The extensive decommissioning of the existing gas plant is more than compensated by the investment in NG-CCS plants that operate on gas but capture most of the emissions [3]. This is consistent with the high import cost of gas in BO_{10} in its cost breakdown. Relatively lower investment in offshore wind generation can be explained by underestimation of wind generation potential as captured by BO_{10} (Fig. 7b). As a result of high decommissioning and low generation expansion, the planning decisions identified by BO_{10} yield higher costs for worst-case projections than BO_2 as shown by the wider error bars in Fig. 7b. Ultimately, BO_{10} only yields slightly lower expected costs than BO_2, and the decision to adopt either investment strategy should reflect the importance a planner places in worst-case costs.

TABLE II: Installed VRE capacity (GW), VRE utilization rate (%), and storage investment ($ billion). Utilization rate is calculated as average hourly generation divided by installed capacity.

	Solar		Offshore		Onshore		Storage
	Cap.	Rate	Cap.	Rate	Cap.	Rate	Inv.
Base. 1	21.4	2.71	7.7	6.00	8.7	8.48	1.00
Base. 2	21.4	2.66	9.0	5.98	8.7	8.49	0.96
Random	21.4	2.74	6.6	6.06	8.7	8.48	1.02
BO_{10}	21.4	2.79	3.3	6.09	8.7	8.50	1.07
BO_2	21.4	2.76	6.5	6.06	8.7	8.51	1.12

Compared to Base. 1 and Base. 2, all search methods invest less in expanding generation capacity but more in energy storage infrastructure (Tab. II). Nevertheless, their resulting investments can substantially reduce costs from load shedding in the power system when compared to Base. 1, and in the case of BO_2, operate with zero load shedding (similar to Base. 2) as a result of including two extreme days (Fig. 8). Interestingly, Tab. II shows that the utilization rate of VRE assets is generally higher for the search approaches than the baselines. In particular, BO_{10}, which weighted solar CFs more heavily in clustering, is able to dispatch more power from the same installed solar capacity through optimizing geographical distribution of plant locations. These findings reflect the cost savings obtained by our BO-assisted approach resulting from higher investment

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

in storage coupled with establishment of relatively fewer VRE assets that can nevertheless be dispatched more effectively through strategic positioning.

V. Concluding Remarks

In this work, we presented a BO-assisted approach for solving large-scale stochastic capacity expansion problems that learns to construct and solve reduced-order models in deployment. Leveraging an established time series aggregation heuristic, we identify lower-cost capacity expansion decisions by optimizing over a continuous, low-dimensional space of representative period clustering hyperparameters. Importantly, we optimize with respect to operational costs incurred by planning decisions on a set of validation scenarios to obtain a less biased expected operational cost estimate. We apply our approach to capacity expansion planning of a coupled power and NG network and show that, when compared to conventional approaches to time series aggregation, our approach is able to reduce load shedding and other operational costs while minimizing investment costs through greater investment in storage and strategic positioning of solar and offshore wind generation assets.

In our discussion of results, we identified two investment strategies, both discovered by BO approaches, that yielded similar costs on average but different costs under worst-case projections. While we focus on minimizing expected costs in this work, future works may leverage recent advances in BO that consider risk measure-based objectives [38] to optimize for worst-case scenarios. Future work is needed to explore learning joint spatio-temporal aggregations of CEPs (following [9]) to more directly consider the effects of geographical variation in VRE availability.

Acknowledgement

The authors would like to acknowledge support from the MIT Energy Initiative Future Energy System Center and MIT Climate Grand Challenge "Preparing for a new world of weather and climate extremes."

References

[1] L. A. Roald, D. Pozo, A. Papavasiliou, D. K. Molzahn, J. Kazempour, and A. Conejo, "Power systems optimization under uncertainty: A review of methods and applications," *Electric Power Systems Research*, vol. 214, p. 108725, 2023.

[2] B. Zhao, A. J. Conejo, and R. Sioshansi, "Coordinated expansion planning of natural gas and electric power systems," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3064–3075, 2017.

[3] R. Khorramfar, D. Mallapragada, and S. Amin, "Electric-gas infrastructure planning for deep decarbonization of energy systems," *Applied Energy*, vol. 354, p. 122176, 2024.

[4] T. Klatzer, U. Bachhiesl, and S. Wogrin, "State-of-the-art expansion planning of integrated power, natural gas, and hydrogen systems," *International Journal of Hydrogen Energy*, vol. 47, no. 47, pp. 20 585–20 603, 2022.

[5] M. Farrokhifar, Y. Nie, and D. Pozo, "Energy systems planning: A survey on models for integrated power and natural gas networks coordination," *Applied Energy*, vol. 262, p. 114567, 2020.

[6] M. M. Frysztacki, J. Hörsch, V. Hagenmeyer, and T. Brown, "The strong effect of network resolution on electricity system models with high shares of wind and solar," *Applied Energy*, vol. 291, p. 116726, 2021.

[7] L. Kotzur, L. Nolting, M. Hoffmann, T. Groß, A. Smolenko, J. Priesmann, H. Büsing, R. Beer, F. Kullmann, B. Singh *et al.*, "A modeler's guide to handle complexity in energy systems optimization," *Advances in Applied Energy*, vol. 4, p. 100063, 2021.

[8] N. Helistö, J. Kiviluoma, H. Holttinen, J. D. Lara, and B.-M. Hodge, "Including operational aspects in the planning of power systems with large amounts of variable generation: A review of modeling approaches," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 8, no. 5, p. e341, 2019.

[9] A. Brenner, R. Khorramfar, and S. Amin, "Learning spatio-temporal aggregations for large-scale capacity expansion problems," in *Proceedings of the ACM/IEEE 14th International Conference on Cyber-Physical Systems (with CPS-IoT Week 2023)*, ser. ICCPS '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 68–77. [Online]. Available: https://doi.org/10.1145/3576841.3585920

[10] H. Teichgraeber and A. R. Brandt, "Time-series aggregation for the optimization of energy systems: Goals, challenges, approaches, and opportunities," *Renewable and Sustainable Energy Reviews*, vol. 157, p. 111984, 2022.

[11] M. Hoffmann, L. Kotzur, D. Stolten, and M. Robinius, "A review on time series aggregation methods for energy system models," *Energies*, vol. 13, no. 3, p. 641, 2020.

[12] H. Teichgraeber and A. R. Brandt, "Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison," *Applied energy*, vol. 239, pp. 1283–1293, 2019.

[13] C. Li, A. J. Conejo, J. D. Siirola, and I. E. Grossmann, "On representative day selection for capacity expansion planning of power systems under extreme operating conditions," *International Journal of Electrical Power & Energy Systems*, vol. 137, p. 107697, 2022.

[14] M. Sun, F. Teng, X. Zhang, G. Strbac, and D. Pudjianto, "Data-driven representative day selection for investment decisions: A cost-oriented approach," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2925–2936, 2019.

[15] D. S. Mallapragada, D. J. Papageorgiou, A. Venkatesh, C. L. Lara, and I. E. Grossmann, "Impact of model resolution on scenario outcomes for electricity sector system expansion," *Energy*, vol. 163, pp. 1231–1244, 2018.

[16] S. Pfenninger, "Dealing with multiple decades of hourly wind and pv time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability," *Applied energy*, vol. 197, pp. 1–13, 2017.

[17] H. Teichgraeber, C. P. Lindenmeyer, N. Baumgärtner, L. Kotzur, D. Stolten, M. Robinius, A. Bardow, and A. R. Brandt, "Extreme events in time series aggregation: A case study for optimal residential energy supply systems," *Applied energy*, vol. 275, p. 115223, 2020.

[18] S. Pineda and J. M. Morales, "Chronological time-period clustering for optimal capacity expansion planning with storage," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 7162–7170, 2018.

[19] V. Oree, S. Z. Sayed Hassen, and P. J. Fleming, "Generation expansion planning optimisation with renewable energy integration: A review," *Renewable and Sustainable Energy Reviews*, vol. 69, pp. 790–803, 2017.

[20] N. E. Koltsaklis and A. S. Dagoumas, "State-of-the-art generation expansion planning: A review," *Applied energy*, vol. 230, pp. 563–589, 2018.

[21] S. Adams, P. Bernard, J. Black, P. Boughan, R. Ethier, E. Foley, A. George, E. Johnson, S. Judd, A. Kleeman, J. Lowe, A. McBride, B. Oberlin, M. Perben, K. Quach, G. Saulmon, D. Schwarting, P. Vijayan, B. Wilson, and M. Winne, "2023 regional system plan," ISO New England Inc., Tech. Rep., 2023.

[22] Y. Bengio, A. Lodi, and A. Prouvost, "Machine learning for combinatorial optimization: a methodological tour d'horizon,"

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

European Journal of Operational Research, vol. 290, no. 2, pp. 405–421, 2021.
[23] T. Han and D. J. Hill, "Learning-based topology optimization of power networks," *IEEE Transactions on Power Systems*, vol. 38, no. 2, pp. 1366–1378, 2023.
[24] E. Prat and S. Chatzivasileiadis, "Learning active constraints to efficiently solve linear bilevel problems: Application to the generator strategic bidding problem," *IEEE Transactions on Power Systems*, vol. 38, no. 3, pp. 2376–2387, 2023.
[25] R. Khorramfar, M. Santoni-Colvin, S. Amin, L. K. Norford, A. Botterud, and D. Mallapragada, "Cost-effective planning of decarbonized power-gas infrastructure to meet the challenges of heating electrification," *arXiv preprint arXiv:2308.16814*, 2023.
[26] K. Poncelet, H. Höschle, E. Delarue, A. Virag, and W. D'haeseleer, "Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 1936–1948, 2016.
[27] A. Almaimouni, A. Ademola-Idowu, J. Nathan Kutz, A. Negash, and D. Kirschen, "Selecting and evaluating representative days for generation expansion planning," in *2018 Power Systems Computation Conference (PSCC)*, 2018, pp. 1–7.
[28] A. Nasri, S. J. Kazempour, A. J. Conejo, and M. Ghandhari, "Network-constrained ac unit commitment under uncertainty: A benders' decomposition approach," *IEEE transactions on power systems*, vol. 31, no. 1, pp. 412–422, 2015.
[29] Y. Zhang, V. Cheng, D. S. Mallapragada, J. Song, and G. He, "A model-adaptive clustering-based time aggregation method for low-carbon energy system optimization," *IEEE Transactions on Sustainable Energy*, vol. 14, no. 1, pp. 55–64, 2022.
[30] L. Kaufman and P. Rousseeuw, *Partitioning Around Medoids (Program PAM)*. John Wiley & Sons, Ltd, 1990, ch. 2, pp. 68–125. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470316801.ch2
[31] A. Almaimouni, A. Ademola-Idowu, J. N. Kutz, A. Negash, and D. Kirschen, "Selecting and evaluating representative days for generation expansion planning," in *2018 Power Systems Computation Conference (PSCC)*. IEEE, 2018, pp. 1–7.
[32] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms," in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012.
[33] P. I. Frazier, "A tutorial on bayesian optimization," 2018.
[34] M. Balandat, B. Karrer, D. R. Jiang, S. Daulton, B. Letham, A. G. Wilson, and E. Bakshy, "BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization," in *Advances in Neural Information Processing Systems 33*, 2020. [Online]. Available: http://arxiv.org/abs/1910.06403
[35] C. E. Rasmussen, C. K. Williams *et al.*, *Gaussian processes for machine learning*. Springer, 2006, vol. 1.
[36] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 1999.
[37] J. Weiss and J. M. Hagerty, "Achieving 80% ghg reduction in new england by 2050," Brattle Group, Inc., Tech. Rep., 2019.
[38] S. Cakmak, R. Astudillo, P. Frazier, and E. Zhou, "Bayesian optimization of risk measures," 2020.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.