Optimizing the Flexible Design of Hybrid Renewable Energy Systems

Abstract—Engineering systems often operate for a long period of time under varying conditions. Designers design the system based on the best available information at the time of the decision. They should also account for future uncertainties in the initial design of the system. The initial design may or may not change as the future evolves and conditions change. The goal of this study is to optimize the design of a hybrid renewable energy system to deliver electricity under highly uncertain demand. This research explores designing the hybrid system while taking into account uncertainties over a long period of time (i.e., 20 years in this study). The objective is to minimize the expected discounted cost of the hybrid renewable energy system during the next 20 years. A design solution may also be flexible, which means that the design can be adapted or modified in the future to meet new scenarios. This paper incorporates flexibility or capacity expansion into engineering design under long-range uncertainty when the objective function is evaluated via a Monte Carlo simulation. The value of expanding capacity is measured by comparing the cost without capacity expansion and cost with capacity expansion.

Index Terms—Engineering Design, Hybrid Renewable Energy System, Flexibility, Capacity Expansion, Monte Carlo Simulation, Bayesian Optimization.

I. Introduction

Traditional engineering design assumes that customers and the public know what they want with a high degree of certainty and that requirements will not change over time. The new system will operate in a stable environment in which the regulations, technologies, demographics, and usage patterns will not change [1]. However, a design may not be successful in the future because the operating conditions or demand for a product may change. For example, an automobile company may design and manufacture cars based on the current fuel price without considering the long-term uncertainty in the fuel price [2]. The demand for this company's products may decrease in the future if its cars are not fuel efficient. Many engineered systems and products may also be used in ways that were not originally intended by the designers. For example, the Global Positioning System (GPS) was originally developed for military use, but a countless number of commercial devices rely on the GPS today. The developers failed to design the original GPS with the capability to extend to these non-military uses [1].

Engineering design can be viewed as a decision-making process, but complexity and uncertainty make decision making for systems design challenging [3]. When an engineering system is designed, there may be large amounts of uncertainty about the future, which makes decisions about how best to design the system difficult [4]. Designers design the system based on the best available information at the time of the decision, and they should also account for future uncertainties in the initial design of the system. The initial design may or may not change as the future evolves and conditions change. For example, power plant designers should consider uncertainties in the price and demand for electricity, but they may also want the ability to change the initial design based on the evolution of price and demand and other conditions (e.g., technology for renewable energy sources) in the future. They may choose to expand capacity under some scenarios and choose not to expand their capacity under other scenarios. Since engineered systems constantly face changes and unpredictability in their environments, these systems should be designed with the capability to respond to future changes [5].

Engineering system design may require optimizing high-dimensional, computationally expensive objective functions under long-range uncertainties. Two examples include the wing design of a high-speed aircraft in the aerospace industry [6] and the crash worthiness design in the automotive industry [7]. Designing with many parameters involves optimizing an objective function in high-dimensional space. The evaluation of the objective function often requires the use of a computationally expensive simulation model. High-dimensional design problems typically have discrete and continuous variables, and there is often uncertainty around these problems. Monte Carlo simulation is often used to explore the design output given design parameters and while considering the uncertainty.

If a designer plans to modify the design, add capacity, or alter the system in the future, then the design decision is a multi-stage decision problem. The problem is frequently stochastic because of uncertainties in the future conditions. Designers should find the optimal design of the system and how that design should be modified under different conditions in the future [8]. This multi-stage design problem can manifest itself in multiple ways.

One such manifestation is determining if and when to expand the a system's capacity [4], [9]–[21]. For example, Wang et al. [16] use game theory and bi-level optimization to find amount to capacity expansion amount under an incomplete information for the competitors in the energy market. Capacity may be discrete units in which case mixed-integer linear programs may be used to determine when and if to expand capacity [17], [19] . Hajipour et al. [18] develop a stochastic program based on a Monte Carlo approach to find the optimal capacity for the components of the Microgrid system with wind farms and energy storage.

Another manifestation of this multi-stage design decision under uncertainty is to produce a flexible design. Flexibility in design enables the designers to review the initial design in the future and provides them with the option to modify the system. A flexible design usually gives designers or the firm the ability to easily modify the design in order to respond to changing circumstances such as increasing or decreasing demand [5]. In recent years, flexibility has been considered in a number of fields from architecture to software design [22]–[26]. Abandoning the system permanently or expanding the capacity to handle more demand in the future are examples of flexibility in engineering design [8].

Studies show that a flexible design can reduce the costs of design by 10-30% in comparison to the standard design [27]. A flexible design may enable a system's owner to more easily expand that system's capacity in the future [4], [20], [21]. Solution methods to this multi-stage stochastic optimization problem include generating dozens of sample paths from the total set of uncertain scenarios [4], meta-heuristics and evolution algorithms [20], and deploying Monte Carlo simulation [21]. These papers [4], [20], [21] assume that the model can be easily evaluated for any input decision. This assumption does not apply to cases where the objective function is estimated via a complex simulation in which it is not computationally efficient to evaluate the model output for every possible input.

In this paper, we propose an algorithm to design a flexible hybrid renewable energy system (HRES) by determining whether or not to expand capacity. The method outlined in this paper differs from the previous literature because our method simulates thousands of scenarios with all of the uncertain parameters to identify the optimal engineering system design. Traditional stochastic programming approaches for capacity expansion and flexibility consider a limited number of scenarios within their optimization algorithm. The multi-stage design algorithm in this paper classifies the scenarios into different categories which allows the optimization algorithm to consider thousands of scenarios and approximates the optimal solution with a smaller dimension.

Since engineered systems, especially large-scale infrastructure, frequently operate for long time, a decision-making framework is needed to incorporate both long-range uncertainties and computationally expensive simulations. The design of the HRES is optimized when the objective function is evaluated using Monte Carlo simulation that incorporates uncertainties over a long period of time (i.e., a 20-year lifespan in this study). Two models are developed to optimize the system design. The first model uses a simulation optimization algorithm that considers 10,000 possible future scenarios, and the design variables are selected that minimize the expected discounted cost. In this model, the initial design of the HRES will be fixed and unchanged during the planning horizon. The second model introduces an algorithm that uses those 10,000 future scenarios to optimize a multi-stage decision model. Applying the algorithm to the HRES enables a decision maker to determine if capacity should be added to the HRES at fixed points in time depending on demand. The algorithm can also be used to determine the value of flexibility if the future decision is a decision about whether or not to modify the design depending on the realization of uncertainty.

The uniqueness of this paper is that it measures the value of expanding capacity in the future or the value of flexibility in complex engineered systems that require computationally expensive simulations to evaluate the objective function. The paper develops a model to optimize the design of such systems under highly uncertain parameters. Our study is the first to use simulation optimization to solve a multi-stage decision problem for an HRES. In this study, a multi-stage algorithm is proposed to find the optimal capacity expansion for the components of the HRES over a 20-year planning horizon. The optimization algorithm measures the value of expanding capacity or flexibility by comparing the value of design with capacity expansion to the value of design without capacity expansion.

The rest of the paper is organized as follows. In section II, the decision-making framework—to include the structure of the HRES, the model for flexibility, and the simulation optimization routine—is discussed in detail. Section III applies the proposed algorithm and the simulation optimization method to a real-world case study. Finally, section IV provides concluding remarks and directions for further research.

II. Decision-Making Framework

A. Hybrid Renewable Energy Systems

The environmental effects of fossil fuel are encouraging greater usage of renewable energy to meet rising energy demand. The high cost of renewable energy technologies is one of the main challenges to greater use of renewable energy sources. To overcome these challenges, renewable energy sources can be integrated to meet the energy demand of a given area. There are different types of HRES such as biomass-wind-fuel cell, photovoltaic-wind-battery, and photovoltaic-wind-battery-fuel cell [28]. The HRES that we consider consists of photovoltaic (solar panel), wind turbine, battery, electrolyzer, hydrogen tank, and fuel cell.

The mathematical model for the HRES comes from [29]–[31]. For instance, Kaviani et al. [29] optimize the design of HRES without considering the uncertainty of wind speed, solar irradiation, and demand. Kaviani et al. [29] state that incorporating the uncertainty of these parameters in the HRES model requires the use of "computationally intensive and time consuming algorithms like Monte Carlo simulations." In those studies, the initial design of the system is fixed during the long-term planning horizon and cannot be changed in the future. This assumption makes their models less applicable to an unpredictable and changing future. Incorporating the flexible design in which the initial design of the HRES can be modified in the future and modeling complex uncertainties represent

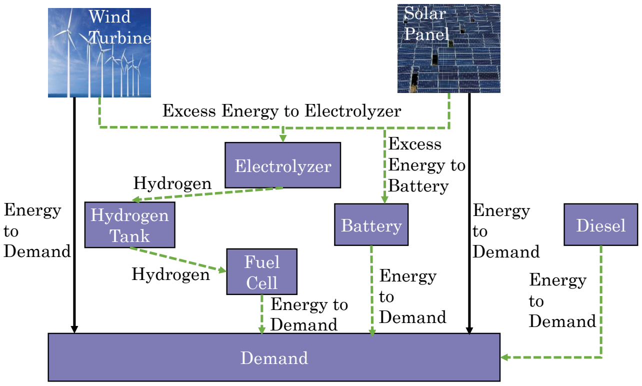


Fig. 1: The energy flow of hybrid renewable energy system

an important advancement in understanding how to best design renewable energy systems.

The solar panel and wind turbine work to generate electricity to satisfy demand. If the generation from solar and wind exceeds demand, then the surplus amount is stored in the battery for the future. If the battery's capacity is exceeded, any excess energy is converted to hydrogen by the electrolyzer and stored in the hydrogen tank. The fuel cell can convert the hydrogen to electricity. The battery and fuel cell will be utilized if the solar and wind generation fail to satisfy demand in a given period. If the wind, solar, and battery sources of energy cannot fulfill demand, the fuel cell can convert the stored hydrogen to electricity. Energy storage systems (i.e., the battery and hydrogen tank) are included in the model to overcome the mismatch between the electricity demand and supply [32]. If the combination of all of these sources cannot satisfy demand, diesel fuel can be used to meet the remaining demand. Figure 1 depicts the energy flow inside the HRES.

1) Solar Panel: The solar photovoltaic (PV) panel is a device that converts the solar irradiation into electricity using solar cells. It can play an important role in generating energy in regions that receive a large amount of sunlight. The hourly output power of the PV panel E_t^{pv} is calculated as:

$$E_t^{pv} = \eta^{pv} \cdot SI_t \cdot A^{pv} \tag{1}$$

where η^{pv} is the efficiency of the PV panel, SI_t indicates the solar irradiation on the surface of the panel at time t, and A^{pv} represents the area of the solar panel.

2) Wind Turbine: Wind turbines convert kinetic energy from the wind into the electrical energy. The output power of the wind turbine E^{wg} is calculated as:

$$E_t^{wg} = \begin{cases} 0.5 C_p \rho A^{wg} u_t^3, & \text{if } u_c < u < u_r \\ 0.5 C_p \rho A^{wg} u_r^3, & \text{if } u_r < u < u_f \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where C_p is the power coefficient, ρ is the air density, and A^{wg} is the area of the rotor, u, u_c, u_r, and u_f are the the wind velocity, cut-in wind velocity, the rated wind velocity, and the cut-off wind speed, respectively [33]. When the wind speed is between the rated wind velocity and cut-off speed then the energy is calculated based on the rated wind velocity.

3) Battery: If the total amount of wind and solar power exceeds demand, then the battery will be charged. Otherwise, the battery is discharged to fulfill the unmet demand. The battery state of charge at time t S_t^{bat} is:

$$S_t^{bat} = \begin{cases} S_{t-1}^{bat} - \eta_{bat}^{d} \times D_t^{bat} / cap_{bat}, & \text{if } \mathrm{D}_t > \mathrm{E}_t^{wg} + \mathrm{E}_t^{pv} \\ \\ S_{t-1}^{bat} + \eta_{bat}^{c} \times E_t^{bat} / cap_{bat}, & \text{if } D_t < E_t^{wg} + E_t^{pv} \end{cases} \tag{3}$$

where η_{bat}^c represents the efficiency of the battery if it is charging, η_{bat}^d when it is discharging, cap_{bat} depicts the capacity of the battery, D_t shows the demand for electricity at time t, and E_t^{bat} is the amount of energy that goes into the battery at time t. The battery is charging when the amount of electricity generated from solar panels and wind turbines is greater than the demand at that period. Otherwise, the battery will be discharged. The battery can be charged until it reaches S_{max}^{bat}, and it can be discharged until it reaches S_{min}^{bat}. Thus, $S_{min}^{bat} < S_t^{bat} < S_{max}^{bat}$.

4) Electrolyzer: The electrolyzer converts electricity into hydrogen [31]. It is directly connected to the hydrogen tank. When the battery is full and a surplus amount of electricity generated by wind and solar exists, the electrolyzer converts the surplus electricity to hydrogen. The amount of energy generated by the electrolyzer to be stored in the hydrogen tank E_{el} is calculated as:

$$E_t^{el} = \begin{cases} \eta^{el} (E_t^{wg} + E_t^{pv} + E_t^{bat} - D_t), & \text{if } D_t > E_t^{wg} + E_t^{pv} + E_t^{bat} \\ & \;\; \& \; S_t^{bat} = S_{max}^{bat} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where η^{el} is the electrolyzer's efficiency in converting electricity to hydrogen.

5) Hydrogen Tank: The hydrogen tank stores the energy produced by electrolyzer. The amount of energy in the hydrogen tank increases when the electrolyzer supplies the hydrogen tank with hydrogen, and the amount of energy in the tank decreases when the fuel cell consumes energy. The amount of energy in the tank at time t E_t^{tank} is:

$$E_t^{tank} = E_{t-1}^{tank} + E_t^{el} - E_t^{fc} \tag{5}$$

where E_t^{fc} is the amount of energy generated by the fuel cell at time t.

6) Fuel Cell: The fuel cell is a device that converts chemical energy into electricity. It can be used as a source of energy to generate electricity. The amount of electricity or energy generated by the fuel cell E_t^{fc} is:

$$E_t^{fc} = \begin{cases} \eta^{fc} \min[D_t - (E_t^{wg} + E_t^{pv} + E_t^{bat}), & \text{if } D_t < E_t^{wg} + E_t^{pv} \\ \;\; \eta^{tank} E_t^{tank})] & \;\; + E_t^{bat} \; \& \; S_t^{bat} = S_{min}^{bat} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where η^{fc} and η^{tank} are the efficiencies of the fuel cell and hydrogen tank, respectively. The fuel cell is utilized when the demand for electricity is greater than the total amount of energy generated by solar panel, wind turbine, and battery and the battery is fully discharged.

7) Diesel Generator: The diesel generator is used as the emergency source of power to satisfy any demand that cannot be met by the renewable sources. The amount of energy needed by the diesel generator at time t E_t^{ds} is calculated as:

$$E_t^{ds} = \begin{cases} \frac{1}{\eta^{ds}}[D_t - (E_t^{wg} + E_t^{pv} \\ +E_t^{bat} + E_t^{fc})], & \text{if } D_t > (E_t^{wg} + E_t^{pv} \\ & +E_t^{bat} + E_t^{fc}) \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

where η^{ds} is the efficiency of the diesel generator.

B. Cost model of HRES

The objective is to minimize the discounted life-cycle cost of the HRES. The cost function consists of four parts: investment, operations and maintenance, replacement, and diesel fuel costs. The total cost of the HRES depends on the size or capacity of each component (i.e., the decision variables). The investment cost INV occurs at the beginning of system operation. The operations and maintenance cost OP occurs in each period during the system's life cycle. Since the system needs to be maintained properly, replacement costs REP occurs when any of the components of the HRES requires replacement. The diesel fuel costs FC represents the cost of purchasing diesel fuel to satisfy demand. If the total amount of supply by HRES at period t is less than the demand, then the shortage amount will be fulfilled by purchasing diesel from the market. The model ignores the power transmission cost between system components and also from the supply to the demand location. The model is a simplified version of the real HRES system where there should be multiple system components at different locations.

The parameters c_i^{inv}, c_i^{om}, and c_i^{rep} are the investment, operations and maintenance, and replacement costs for the ith design component. The $I = 6$ design components of the HRES are: the PV panel, the wind turbine, the battery, the electrolyzer, the hydrogen tank, and the fuel cell. Each component has an energy capacity cap_i. The number of times the ith component will be replaced is R_i. L_i is lifetime of component i. The planning-time horizon has T total periods, and λ represents the interest rate. The parameter c_{ds} is the diesel cost. Eqs. (8-12) provide the formula for calculating the life-cycle cost and its four cost components. Eq. (8) shows the total cost of the system when operating for the T periods. Eq. (9) shows the investment cost. Eq. (10) displays the annual cost of operations and maintenance which is converted to the present value. Eq. (11) shows the present value of the cost to replace the system's components at the end of their lifetime. Eq. (12) calculates the present value of the fuel cost when the renewable energy sources cannot fulfill demand.

$$\text{cost} = \text{INV} + \text{OP} + \text{REP} + \text{FC} \tag{8}$$

$$\text{INV} = \sum_{i=1}^{I} c_i^{inv} \text{cap}_i \tag{9}$$

$$\text{OP} = \sum_{i=1}^{I} c_i^{om} \frac{(1+\lambda)^T - 1}{\lambda(1+\lambda)^T} \text{cap}_i \tag{10}$$

$$\text{REP} = \sum_{i=1}^{I} \sum_{r=1}^{R_i} c_i^{rep} \frac{1}{(1+\lambda)^{L_i \times r}} \text{cap}_i \tag{11}$$

$$\text{FC} = c_{ds} \sum_{t=1}^{T} \frac{1}{(1+\lambda)^t} \text{E}_{ds,t} \tag{12}$$

Complexity and dynamics inside the hybrid renewable energy system require the use of Monte Carlo simulation to calculate the cost function. Monte Carlo simulation is used to propagate the parameter uncertainties to the uncertainty in the life-cycle cost. The expected cost of design is calculated as the average after N different simulations.

The decision variables for designing the HRES are the capacity of each component, cap_i. The energy generated by each component at time t, E_t^{pv}, E_t^{wg}, E_t^{bat}, E_t^{el}, E_t^{tank}, and E_t^{fc}, must not exceed the chosen capacity for each component cap_i. Each component also has a maximum capacity, cap_i^{max}, so that

$$E_t^i \leq cap_i \;\; \forall i = 1, \ldots, I, \;\; t = 1, \ldots, T \tag{13}$$

and

$$0 \leq \text{cap}_i \leq \text{cap}_i^{max} \;\; \forall i = 1, \ldots, I \tag{14}$$

where E_t^i represents the energy generated by component i at time t.

The design decision maker should choose to minimize the expected discounted life-cycle cost of the HRES by choosing the capacity of each component and ensuring the constraints in Eqs. (13) and (14) be satisfied.

C. Multi-Stage Decision-Making Model

Engineering design may involve multiple decisions over time. One type of multi-stage design is a flexible design, and a flexible design will allow for different designs, each of which depends on the realization of individual scenarios. In the static design without flexibility, the decision maker would design the HRES based on all of the future demand and cost simulations from the current time to the end of planning horizon. However, in the design with flexibility, the designers have the option to decide whether or not to expand the capacities of the HRES if it is needed to generate more electricity to meet increasing future demand.

The procedure of multi-stage design when the objective function should be evaluated with the computationally expensive simulations cannot be done rolling back from the end of simulation and exercising the options at each period of time. In this case, the model should be optimized for each of N simulations, which is usually large (10,000 simulations or more), and for each period in the planning time horizon, T. Solving the optimization model will suffer from the curse of dimensionality. In the proposed algorithm for design for flexibility (see algorithm 1), we discretize the continuous HRES problem and consider a multi-stage decision making process for flexible design. The decision maker has the option to expand the capacity of each component during each review of system. For example, the decision to expand the capacity of each component of the HRES could occur every 5 years.

The proposed algorithm starts by optimizing the model (i.e., Eqs. (1-14)) during the first T_1 periods (e.g., $T_1 = 10$ years) by taking into account all of the N future scenarios from time 0 to T_1 (lines 2-4 in algorithm 1). The initial optimal design will be used as an input into the design modification stages. At time T_1, the first decision for the capacity expansion will be made. The future scenarios from the first design modification period to the next review of the system (periods T_1 to T_2) are divided into K_2 different categories. For example, if the future scenarios consist of demand for electricity, the future demand could be categorized into low, medium, and high demand during periods T_1 to T_2. Given the initial optimal design, the algorithm optimizes the additional capacity by minimizing the expected cost for each of the K_2 categories from T_1 to T_2. At the end of stage 2, there will be K_2 different designs and expected costs (lines 5-10).

At stage 3, given the initial optimal design and each of the K_2 different additional categories, the expected cost will be minimized by optimizing Eqs. (1-14). The future scenarios from the second design modification period to the next review of the system (periods T_2 to T_3) are divided into K_3 different categories (e.g., low, medium, and high demand). Given the initial optimal design and each optimal expansion amounts of stage 2, the algorithm optimizes the additional capacities by minimizing the expected cost for each of the K_3 categories from periods T_2 to T_3. In stage 3, there will be $K_2 K_3$ different additional capacities. There will be K_3 optimal additional capacities for each of K_2 expansion amounts at stage 2 (lines 11-19). This process will continue S times where S is the number of modification stages.

At stage $S+1$, given the initial optimal design and all additional capacities in the previous S design stages, the expected cost will be minimized by optimizing Eqs. (1-14). The demand from stage $S+1$ of the design modification period to the end of planning horizon is divided into K_{S+1} different regions. Given the initial optimal design and each of the optimal expansion amounts of all previous S stages, the algorithm optimizes the additional capacities by minimizing the expected cost for each of the K_{S+1} categories from T_S to T. In stage $S+1$, there will be $\prod_{s=1}^{S} K_{s+1}$ optimal additional capacities for each K_{S+1} expansion amounts at stage $S+1$ (lines 21-31 of algorithm 1). At each decision-making stage, the capacities of the components can be increased to fulfill demand for the upcoming planning periods.

The total expected cost of the design ECF is the sum of the initial expected cost $E[cost^1]$ and average capacity expansion costs from stage 1 to stage $S+1$,

Algorithm 1 Multi-stage decision-making model

```
1:  Inputs: N future scenarios of demand for t = (0,T), S number of design stages where each design stage occurs at discrete times 0, T_1, T_2, ..., T.
2:  Stage 1
3:  Solve Eqs. (1-14) and find optimal design for N scenarios from t = (0 : T_1)
4:  Outputs: E[cost^1] and cap^1
5:  Stage 2
6:  Inputs: N future demand scenarios from time T_1 to T_2 described as D^{T_1:T_2} and cap^1
7:  for k_2 ← 1 to K_2 do
8:      Solve Eqs. (1-14) and find the optimal additional capacity |D = *D_{k_2}^{T_1:T_2}
9:  end for
10: Outputs: K_2 different E[cost^{1,k_2}] and cap^{1,k_2}
11: Stage 3
12: Inputs: N future demand scenarios from time T_2 to T_3 described as D^{T_2:T_3}, cap^1 , and K_2 different cap^{1,k_2}
13: for k_3 ← 1 to K_3 do
14:     for k_2 ← 1 to K_2 do
15:         Solve Eqs. (1-14) and find optimal additional capacity |D = D_{k_3}^{T_2:T_3}
16:     end for
17: end for
18: Outputs: K_2 K_3 different E[cost^{1,k_2,k_3}] and cap^{1,k_2,k_3}
19: ⋮
20: Stage S+1
21: Inputs: N future demand scenarios from time T_S to T described as D^{T_S:T}, cap^1, K_2 different cap^{1,k2}, K_2 K_3 different cap^{1,k_2,k_3}, ..., K_2 K_3 ... K_S different cap^{1,k_2,k_3,...,k_S}
22: for k_{s+1} ← 1 to K_{S+1} do
23:     for k_s ← 1 to K_S do
24:         ⋱
25:         for k_2 ← 1 to K_2 do
26:             Solve Eqs. (1-14) and find optimal additional capacity |D = D_{k_s}^{T_s:T}
27:         end for
28:     end for
29: end for
30: Outputs: ∏_{s=1}^{S} K_{s+1} different E[cost^{1,k_2,k_3,...,k_{S+1}}] and cap^{1,k_2,k_3,...,k_{S+1}}
31: *D_{k_2}^{T_1:T_2} is demand (D) from T_1 to T_2 in which the value of D at time T_1 is within the region k_2
```

which are discounted by the interest rate λ. Since the number of different decisions to expand capacities in stage s is $\prod_{s'=1}^{s} K_{s'+1}$, the expected cost in stage s for category $k_s = 1, \ldots, K_s$ is denoted as $E[cost^{1,k_2,k_3,\ldots,k_s}]$. The total expected cost of the design is shown in Eq. (15). Eq. (15) assumes that each of the k_s categories in design stage s occurs with equal probability. If one category is more likely than another category, Eq. (15) can be modified to include the specific probability of each category.

$$\begin{aligned} ECF =& E[cost^1] + \frac{1}{K_2}\frac{1}{(1+\lambda)^{T_1}}\sum_{k_2=1}^{K_2} E[cost^{1,k_2}] \\ &+ \frac{1}{K_2K_3}\frac{1}{(1+\lambda)^{T_2}}\sum_{k_2=1}^{K_2}\sum_{k_3=1}^{K_3} E[cost^{1,k_2,k_3}] + \\ &\ldots + \frac{1}{\prod_{s=1}^{S} K_{s+1}}\frac{1}{(1+\lambda)^{T_s}}\sum_{k_2=1}^{K_2}\sum_{k_3=1}^{K_3}\ldots\sum_{k_{S+1}=1}^{K_{S+1}} E[cost^{1,k_2,\ldots,k_{S+1}}] \end{aligned} \tag{15}$$

The proposed algorithm for multi-stage decision making is able to find the optimal design for an arbitrary number of categories. However, increasing the number of categories and design stages may make solving the decision impractical. A four-stage problem with 10 categories at each stage requires solving 1+10+100+1000 = 1111 optimization problems. However, a four-stage problem with three categories only requires solving 1+3+9+27 = 40 optimization problems, which is much more practical to solve.

D. Optimization Algorithm

Black-box functions (e.g., f) provide system output for specified values of system inputs, and there typically exists little information about the properties of f . The optimization of a black-box system is referred to as black-box optimization. Black-box optimization algorithms optimize the objective function f through a query of the value of $f(x)$ for a point x, but they cannot make any assumptions on the analytic form of f [34].

Bayesian optimization is a method to utilize the input-output relationship of the black-box systems. Many design decisions are evaluated using complex simulation models, and having tools that can optimize the design using such simulations is important for many practical engineering problems. To optimize such models, the key challenge is that the designer has little knowledge about how the objective function changes with respect to changes in the design decision variables. Such difficulty makes many traditional algorithms that require either first-order (i.e., gradient) or second-order (i.e., Hessian) information invalid. Bayesian optimization algorithms can effectively solve problems that seek to integrate optimization into simulation analysis [35]. Further details about Bayesian optimization is discussed in [36].

The decision variable $cap = [cap_{pv}, cap_{wg}, cap_{bat}, cap_{el}, cap_{tank}, cap_{fc}]$ is a vector composed of the capacities of all of the components of the HRES (i.e., solar panel, wind turbine, battery, electrolyzer, fuel tank, and fuel cell). The Bayesian optimization algorithm evaluates the objective function without using any first-order or second-order information. The algorithm constructs a surrogate model for the objective function and exploits the surrogate model to find the next evaluation points in the feasible solution space [37]. A surrogate model is used since the objective function cannot be directly calculated. The surrogate model helps us to estimate the black-box function by constructing an approximate model. The prior and acquisition functions are the two major ingredients of the Bayesian optimization algorithm. We choose the Gaussian process as the prior over the objective function depicted in Eq. (8) as it is a powerful prior distribution for functions [38]. The objective function is approximated with a multivariate normal probability distribution function. The predictive distribution for the unobserved values of the objective function (f^{new}) has the following form:

$$\begin{aligned} &p(f^{new}|\mathrm{cap}^{new}, f(\mathrm{cap}_{1:n}), \mathrm{cap}_{1:n}) \\ &= N(f^*|\mu(\mathrm{cap}^*|\mathrm{cap}_{1:n}), \sigma^2(\mathrm{cap}^*|\mathrm{cap}_{1:n})) \end{aligned} \tag{16}$$

$$\begin{aligned} &\mu(\mathrm{cap}^*|\mathrm{cap}_{1:n}) = \\ &Cov(\mathrm{cap}^*, \mathrm{cap}_{1:n}) Cov(\mathrm{cap}_{1:n}, \mathrm{cap}_{1:n})^{-1} f(\mathrm{cap}_{1:n}) \end{aligned} \tag{17}$$

$$\begin{aligned} &\sigma^2(\mathrm{cap}^*|\mathrm{cap}_{1:n}) = Cov(\mathrm{cap}^*, \mathrm{cap}^*) - \\ &Cov(\mathrm{cap}^*, \mathrm{cap}_{1:n}) Cov(\mathrm{cap}_{1:n}, \mathrm{cap}_{1:n})^{-1} Cov(\mathrm{cap}^*, \mathrm{cap}_{1:n})^T \end{aligned} \tag{18}$$

where $\mathrm{cap}_{1:n}$ is the n previously evaluated capacities used to predict the next point, $Cov(\mathrm{cap}_{1:n}, \mathrm{cap}_{1:n})$ is the covariance matrix for the n design alternatives that are simulated to estimate the objective function, $Cov(\mathrm{cap}^*, \mathrm{cap}_{1:n})$ is the covariance of the n design variables and the new capacity cap^* to be evaluated, $\mu(\mathrm{cap}^*|\mathrm{cap}_{1:n})$ is the posterior predictive mean, and $\sigma^2(\mathrm{cap}^*|\mathrm{cap}_{1:n})$ is the posterior predictive variance. The Gaussian radial basis function kernel is used to calculate the covariance between two capacities cap and $\mathrm{cap}^{'}$:

$$Cov(\mathrm{cap}, \mathrm{cap}^{'}) = \exp\left(\frac{||\mathrm{cap} - \mathrm{cap}^{'}||^2}{2\gamma^2}\right) \tag{19}$$

where γ is a free parameter which will be tuned in the optimization. The current best capacity that results in the smallest expected cost based on n evaluated design alternatives is denoted as $\mathrm{cap}_{best} = \mathrm{argmin}_{\mathrm{cap}_{1:n}} f(\mathrm{cap}_{1:n})$. The objective function evaluation (i.e., expected discounted cost) corresponding to cap_{best} is $f(\mathrm{cap}_{best})$. The next alternative to sample in the simulation is found by maximizing the acquisition function which is the expected improvement EI over the current best [39].

$$\begin{aligned} EI(\mathrm{cap}^*|\mathrm{cap}_{1:n}) = (\mu(\mathrm{cap}^*|\mathrm{cap}_{1:n}) - f(\mathrm{cap}_{best}))\Phi(Z) \\ + \sigma(\mathrm{cap}^*|\mathrm{cap}_{1:n})\Phi(Z) \end{aligned} \tag{20}$$

where $Z = \frac{\mu(\mathrm{cap}^*|\mathrm{cap}_{1:n}) - f(\mathrm{cap}_{best})}{\sigma(\mathrm{cap}^*|\mathrm{cap}_{1:n})}$ and $\Phi(\cdot)$ denotes the cumulative standard normal distribution function.

The Bayesian optimization algorithm begins with a selection of acquisition function and prior over the objective function Algorithm 2 shows the main steps of Bayesian optimization. The prior has a multivariate normal distribution with six decision variables, and the acquisition function is calculated based on Eq. (20). In step 2, all 10,000 observations from the Monte Carlo simulation are used to calculate the objective function (expected cost).

The decision variables that maximize the expected improvement *EI* become the next alternative to act as input into the simulation. which then will be used as a prior to calculate the posterior. This procedure iterates until there is not enough improvement in the objective function (i.e, termination criteria). We use the Random Embedding Bayesian Optimization (REMBO) developed by Wang et al. [40] to implement the Bayesian optimization algorithm in MATLAB software [41].

Algorithm 2 Implementation procedure of Bayesian optimization.

1: Choose an appropriate acquisition function and prior over the objective function
2: Find the posterior over the objective function given some observations
3: Use the posterior and an appropriate acquisition function to identify the next evaluation points
4: Update design alternatives and their corresponding expected costs functions set

III. Application

In this section, the design of HRES is optimized to deliver electricity for the state of California under highly uncertain demand. The planning horizon is 20 years (from 2017 to 2036) and each period is 1 month. Demand for electricity, solar irradiation, and wind velocity are the uncertain parameters. It is assumed that the hourly solar irradiation is normally distributed with a mean of 0.5 kwh/m^2 and a standard deviation of 0.1. The wind velocity is normally distributed with a mean of 5 m/h and standard deviation of 1. Random numbers from the distribution of the solar irradiation and wind speed are selected and included in the simulation of the dynamics of the system. The interest rate, λ, is 2% per year. It is assumed that the efficiency of the PV panel is $\eta^{pv} = 0.15$. The power coefficient of wind turbine is $C_p = 0.59$, and air density is $\rho = 1.225$. The cut-in wind, rated wind velocities, and cut-off wind speed are $u_c = 3.5$ m/h, $u_r = 14$ m/h, $u_f = 25$ m/h, respectively. The application assumes that the maximum and minimum state of battery are $S^{bat}_{max} = 1$ and $S^{bat}_{min} = 0.2$. The efficiency of the electrolyzer, fuel cell, hydrogen tank, and diesel generator are $\eta^{el} = 0.6$, $\eta^{fc} = 0.6$, $\eta^{tank} = 0.95$ and $\eta^{ds} = 0.95$, respectively.

The investment and replacement cost parameters follow triangular distributions. Table I shows the investment, maintenance, and replacement cost parameters and the lifetime of the components for the HRES. The cost function parameters come from Sharafi et. al [30]. In order to build a close to 100% renewable energy system, the diesel cost (c_{ds}) has a high fixed penalty cost of 8 million dollars per 1 MW.

A simulation-optimization approach is used in order to find the optimal design for the HRES. As described in Section II, each component in the HRES has a relatively complicated dynamic, and calculating the system costs depends on several uncertainties including demand, solar irradiation, wind speed, and component costs and lifetimes. Thus, the HRES model is evaluated with a Monte Carlo simulation. The output of the simulation determines the objective function values that enter into the Bayesian optimization algorithm to identify the optimal design. The goal is to minimize the expected cost function that results from the interaction among the components of the HRES as both supply and demand fluctuate over a time span of 20 years. The Monte Carlo simulation enables the decision maker to analyze the stochastic behavior of the HRES.

A. Demand Forecast

Renewable energy systems are designed for long-term usage. Therefore, it is necessary to establish these sources of electricity generation considering possible future scenarios which will substantially impact the design. Uncertainty in the future demand will impact the design of the HRES, and forecasting demand is required for the control of power systems [42]. Good demand forecasting is an essential prerequisite of an energy system study for the capacity expansion planning. Since future demand will be uncertain, incorporating that uncertainty into the forecasting model will help the system manage load efficiently [43], [44]. One of the goals of this paper is to investigate how the long-range demand uncertainty impacts the design of the HRES.

Demand for electricity is serially autocorrelated, which suggests that a time series model may be appropriate. We choose the Auto Regressive Integrated Moving Average (ARIMA) [45]. ARIMA is a well-known model for time series analysis. In time series forecasting, the goal is to predict a series that typically is not deterministic but rather contains a random component. The ARIMA model is a linear combination of past observed data points (here demand for electricity) and errors to produce a forecast [46], [47]. This method is widely used in the literature to forecast demand for electricity [47]–[53]. The general form of ARIMA(p,d,q) forecasts the future based on the following:

$$(1 - \sum_{i=1}^{p} \wp_i \Im_i)(1-\Im)^d D_t = (1 + \sum_{i=1}^{q} \theta_i \Im^i)\varepsilon_t \qquad (21)$$

where p is the lag order, d is the degree of differencing, q is the order of the moving average, $\Im$ is the lag operator, $\wp$ is the autoregressive operator represented as a polynomial in the lag operator, and θ is the moving-average operator represented as a polynomial in the lag operator. In this case study, $p = 1$, $d = 0$, and $q = 12$. The forecast variable D_t is the demand for electricity at time t, and ε_t is a normally distributed error. The *arima* function in MATLAB software [41] is used to generate the ARIMA model for electricity demand using historical data of monthly electricity demand for California from 2001 to 2016 [54].

Since many different events could occur over 20 years that would have a large impact on electricity demand in

TABLE I: The cost (in millions of $ per 1 MW) and lifetime parameters of the components of the HRES [30]

Component	Lifetime (years)			c_i^{inv}			c_i^{rep}			c_i^{om}
	lower limit	mode	upper limit	lower limit	mode	upper limit	lower limit	mode	upper limit	
Wind	10	25	30	5	7	9	5	6	7.5	20
Solar	10	20	25	1.5	2.5	3	1.2	2	2.5	75
Battery	1	5	7	1.5	2	2.2	1.3	1.5	2.1	20
Electrolyzer	5	10	13	1	2	3	0.9	1.5	2	25
Fuel Tank	10	20	25	0.8	1.3	1.5	0.8	1.2	1.3	15
Fuel Cell	0.7	1.7	2.7	1	3	4	1.9	2.5	3	172

California, we consider a large amount of uncertainty in the future electricity demand in California. The error term ε_t follows a Gaussian distribution with the mean of 0, and variance of 20,000.

Monte Carlo simulation is used to generate 10,000 simulated paths of demand by sampling from ε_t for $t = 1, 2, \ldots, 240$ months and applying the ARIMA model on Eq. (21). Figure 2 shows the historical monthly demand and the generated scenarios for electricity demand for 20 years. For the first five years of planning (2017-2021), the simulated demand scenarios exhibit less uncertainty where 90% of the variation around demand mean lies within the interval of [18.1, 21.5] million Mwh. From 2021 to the end of the planning horizon (i.e., 2036), the uncertainty in demand accumulates which leads to a wide variety of demand paths in the future. Although the demand uncertainty is very large at the end of planning horizon, 90% of variation around demand mean at the end of 2036 is contained within the interval of [16.7, 39.7] million Mwh. Modeling the demand with large uncertainty that covers 20 years into the future enables designers to consider extreme demand scenarios, which may not be very likely but may still need to be considered in policy planning. Such a large uncertainty can help motivate the need for flexibility or capacity expansion in capacity planning.

B. Design without capacity expansion

In the Monte Carlo simulation of the time series model, 10,000 scenarios are generated for modeling the demand for the next 20 years. The model is solved with 10,000 samples from uncertain parameters (demand, cost coefficients, wind speed, and solar irradiation) for 240 months (2017-2036). The simulation optimization algorithm uses Eqs. (1-14) to identify the design alternative that minimizes the expected discounted cost. In design without capacity expansion, the system is designed at the beginning of the system operation in 2017 and will not be modified in the future.

Table II shows the optimal results for design without capacity expansion. The optimal capacities represent the nominal capacity of the components. The actual capacity of each component is the nominal capacity multiplied by the efficiency. For instance, the solar panel is built with 392 GW nominal capacity; however, the actual capacity is $0.15 \times 392 = 58.8$ GW.

The results show that 78% of the demand during the 10,000 simulations from 2017-2036 are fulfilled with solar and wind generation. These two components supply the majority of demand. Since the amount of energy generated by these two sources exceed the demand for many time periods, the surplus amount of energy will be conserved in the battery for future use. When the battery is full, the electrolyzer converts the energy into hydrogen, which is stored in the hydrogen tank and will be converted to electricity by the fuel cell. For instance, based on one of the 10,000 simulations (e.g., Figure 3a), the amount of energy generated by wind and solar almost always exceeds demand during the entire 20 years. There are a few instances in that simulation when solar and wind generation is low that energy from the battery is used.

TABLE II: The optimal design of the HRES for design without capacity expansion

Plant	Optimal Capacity (GW)	Percentage (%)
Solar panel	392	56
Wind turbine	146	22
Battery	89	17
Electrolyzer	1041	-
Hydrogen tank	3221	-
Fuel cell	138	4
Diesel	-	<1

The results show that the battery and fuel cell satisfy 17% and 4% of the demand, respectively. The HRES requires diesel to meet approximately 1% of the demand. The optimal design of the HRES fulfills more than 99% of the electricity with the renewable sources. The expected discounted cost of the design is $40.66 trillion, which includes a $9.56 trillion investment cost, $21.66 trillion in operations and maintenance costs, $9.4 trillion in replacement costs, $40 billion in fuel costs. The large life-cycle cost is due to needing to build huge components for the HRES system. Fulfilling the total demand for electricity in the state of California with HRES requires a large amount of investment and maintenance cost. A study shows that tens of trillions of dollars are needed to achieve California's 80% greenhouse gas reduction target in 2050 by constructing renewable energy systems [55].

Figures 3a, 3b, and 3c show three random simulations to illustrate how demand is fulfilled with different sources of energy in different simulations. In some cases (Figures 3a and 3c), the electricity generated by solar and wind exceeds the demand most of the time. The battery stores the surplus electricity and is used to satisfy demand a small proportion of the time in these two simulations. Figure 3b illustrates that in some scenarios solar and wind generation cannot fulfill the increasing demand in

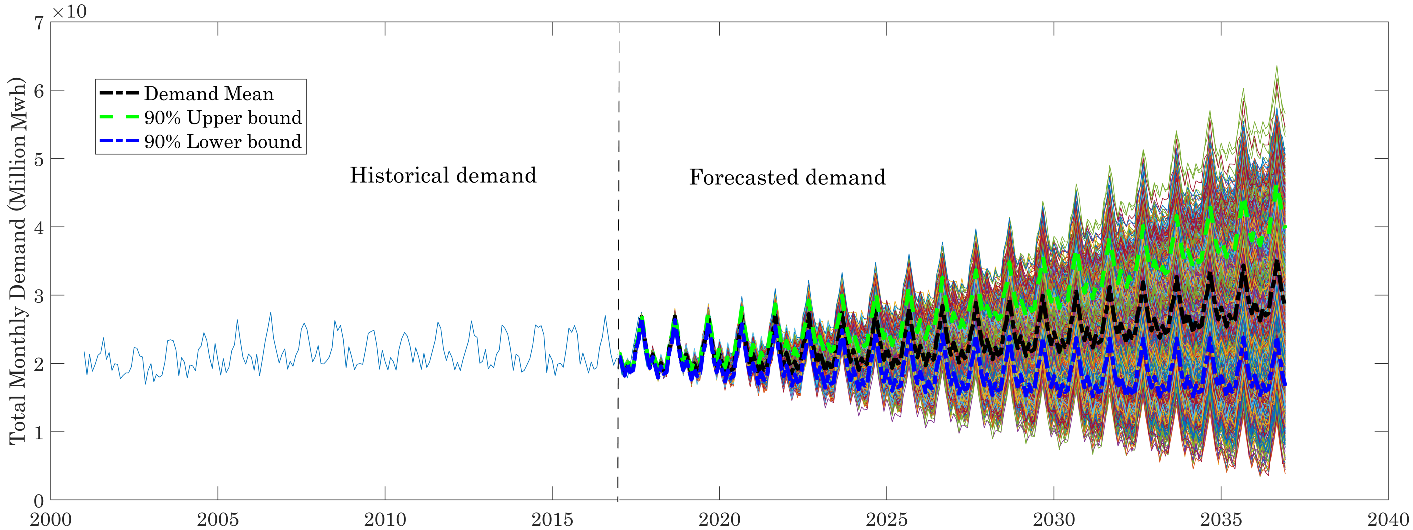


Fig. 2: Simulation of electricity demand for California, 2017-2036

the future. Although the battery helps to satisfy some of the demand, most of the surplus energy generated from years 2017-2031 is stored in the hydrogen tank. The fuel cell converts the hydrogen to electricity and satisfies much of the shortfall in demand from 2031-2034. In 2035, the hydrogen tank is depleted and the fuel cell is unable to provide any more electricity. Diesel fuel is required it meet demand in 2035-2036.

To evaluate the effect of the number of simulation replications, the expected cost of the optimal design is calculated for different numbers of replications. Ten thousand replications result in a half-width for a 95% confidence interval that is approximately 0.5% of the expected cost. Including more than 10,000 replications would drastically increase the computation time without substantially reducing the variation in the objective function. Ten thousand replications are sufficient to estimate the objective function in the Monte Carlo simulation.

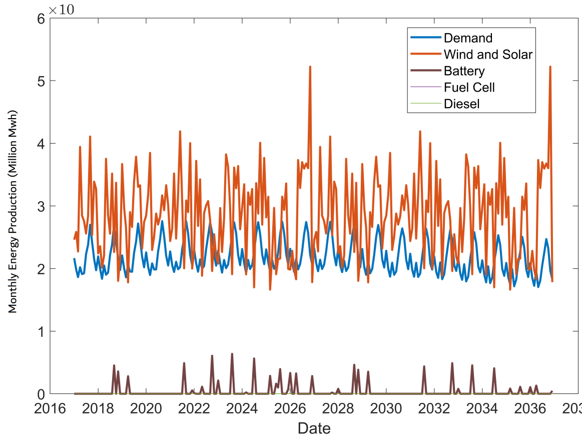
Design decision making involves making trade-offs among many design variables and attributes. Determining how to make those trade-offs may be difficult in complex engineered systems [56]. Optimization helps the decision maker to find the optimal design that offers the best trade-off among all possible design alternatives. Each design alternative can lead to a different expected cost. To better visualize the trade space, parallel coordinate plots can show design alternatives with respect to different attributes and design variables [56]. The set of feasible alternatives from the objective space is projected on the parallel coordinate plot in Figure 4. The plot illustrates how the 50 different combinations of design variables impact the expected cost. The 50 combinations are the 50 best results of the optimization algorithm. It shows the sensitivity of the objective function to each decision variable. The expected cost is calculated based on the actual simulation. The optimal design is represented by the black line. The red line is the second best solution with an expected cost of \$46.25 trillion among those 50 best solutions in the model. This figure shows the complexity of the HRES model. If the value for each variable changes, a large difference in the expected cost may be seen. For instance, if the solar panel's capacity is 176 GW and the wind turbine's capacity is 312 GW, then more capacity is needed for the fuel cell (691 GW) with an expected cost of \$85.27 trillion.

C. Design with capacity expansion

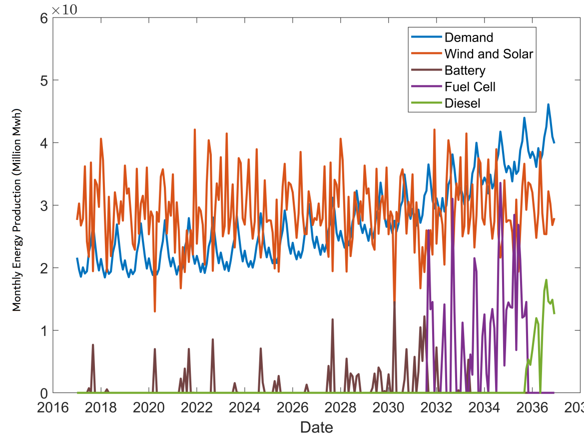
The design with capacity expansion strategy requires a smaller initial investment than the design without capacity expansion. This strategy defers additional costs to the future and takes advantage of the time value of money [8]. A flexible design usually gives designers or the company the ability to easily modify the design in order to respond to changing circumstances. In this section, the value of capacity expansion will be measured by comparing the expected discounted cost of designing with capacity expansion and the expected discounted cost of designing without capacity expansion. Algorithm 1 identifies a flexible design alternative by factoring in the cost of modifying the design in the future.

1) Case 1: One stage capacity expansion modeling: Two different cases for this design are developed. In the first case, one additional stage for the design modification is considered and $S = 1$ and $T_1 = 10$. In stage 1, the initial design and expected discounted cost considers the uncertain demand profiles for 2017-2026. The results of this first stage decision making show that the initial optimal design of the components of the HRES have less capacity compared to the optimal solution in the design without capacity expansion model (see Table III).

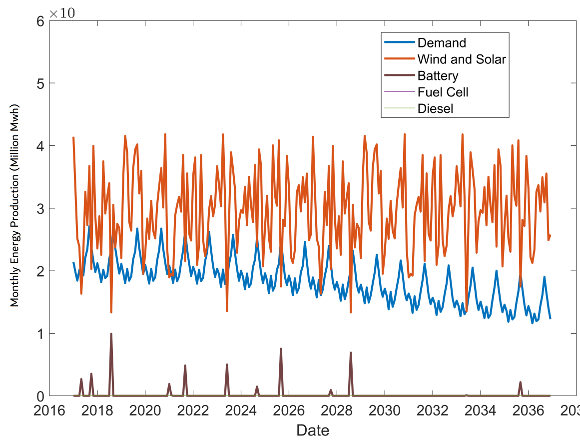
The initial optimal design from 2017-2026 serves as an input to decision making in stage 2, which covers 2027-2036. The stage 2 decision making considers three different demand profiles (i.e., $K_2 = 3$). If demand in month 120,



(a) Demand fulfillment in simulation with constant demand



(b) Demand fulfillment in simulation with raising demand



(c) Demand fulfillment in simulation with decreasing demand

Fig. 3: Demand fulfillment for three random demand scenarios

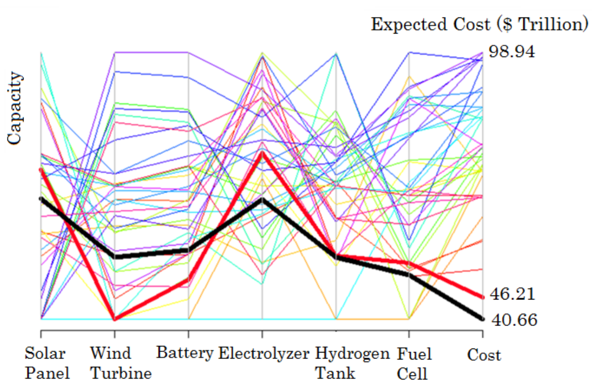


Fig. 4: Parallel coordinates plot for design without capacity expansion

the last month in year 2026, is less than 20.2 million Mwh, the demand is considered low. If demand in month 120 is greater than 20.2 million Mwh and less than 22.7 million Mwh, the demand is considered medium. If demand in month 120 is greater than 22.7 million Mwh, the demand is considered high. Each of these demand profiles occur in 33% of the simulations. Given the initial optimal design, the Bayesian optimization determines if additional capacity for the HRES should be constructed if demand is low, if demand is medium, and if demand is high. Stage 2 contains three different sets of design variables and three different expected discounted costs, one for each demand profile. The average expansion costs are calculated as the expected cost of additional capacity at stage 2. The total expected cost of this design is calculated using Eq. (15).

Table III shows the optimal design for the HRES with capacity expansion with a possible design modification in 2027. Although the wind turbines should initially be constructed with a relatively small capacity in order to avoid unnecessary capacity during the first 10 years, the decision maker should expand the capacity of the wind turbines and the battery in the high or medium demand profiles. Since the wind turbines should be expanded, more battery capacity is needed to store the surplus amounts of energy. The results also show that in the low demand profile, the capacity of the HRES should not be expanded. Diesel should be purchased from the market to supply electricity if demand increases in years 2027-2036. Table III shows that the optimization algorithm chooses the solar panel with a higher initial capacity compared to the wind turbine; however, in the second stage, the algorithm finds that the wind turbine should be expanded and the solar panel should be operating for the next 10 years with the initial capacity.

The results show that the initial expected cost in stage 1 (\$20.55 trillion) is almost half of the expected cost of design without capacity expansion (\$40.66 trillion). However, this initial low capacity design will be expanded in the medium and high demand scenarios which increases the expected cost beyond just the stage 1 cost. The total expected cost of design with capacity expansion for case 1 is \$27.22 trillion. The design with capacity expansion enables the system to defer the additional cost of investment and replacement to the future, avoid the operation and maintenance cost of the full deployment for the first 10 years of operation, and take advantage of the time value of money. The value of capacity expansion is the difference between the expected cost of design with capacity expansion and the expected design without capacity expansion. In this case, the value of capacity expansion is $\$13.44 = \$40.66 - \$27.22$ trillion, a 33% percent reduction in cost.

2) Case 2: Two stage capacity expansion modeling: The second case assumes that the design can be modified twice, once in year 2027 and once in year 2032 ($T_1 = 10$, $T_2 = 15$). Similar to the first case, there are three different regions in each stage of the design modification (i.e., $K_2 = K_3 = 3$). Based on Algorithm 1, at stage 2, three models should be solved for 3 different ranges of demand for low, medium, and high demand profiles. The definitions of the demand profiles in 2027 are equivalent to those in case 1.

The initial optimal design from 2017-2026 serves as an input to decision making in stage 2, which covers 2027-2031. Given the initial optimal design, the Bayesian optimization determines if additional capacity for the HRES should be constructed if demand is low, if demand is medium, and if demand is high. Stage 2 contains three different sets of design variables and three different expected discounted costs, one for each demand profile. The stage 3 decision making also considers low, medium, and high demand profiles. Low demand in month 180 is less than 22.0 million Mwh, medium demand is between 22.0 and 26.3 million Mwh, and high demand is greater than 26.3 million Mwh.

The initial optimal design from 2017-2026 and optimal additional capacities for the three different demand profiles in stage 2 serve as the inputs to decision making in stage 3, which covers 2032-2036. Given the initial optimal design, the Bayesian optimization determines if additional capacity for the HRES should be constructed at 2032 for the following 9 scenarios: high demand in both stages 2 and 3, high demand at stage 3 and medium demand at stage 2, high demand at stage 3 and low demand at stage 2, medium demand at stage 3 and high demand at stage 2, medium demand in both stages 2 and 3, medium demand at stage 3 and low demand at stage 2, low demand at stage 3 and high demand at stage 2, low demand at stage 3 and medium demand at stage 2, and low demand in both stages 2 and 3.

The average expansion cost is calculated as the expected cost of additional capacity at stage 2 plus the expected cost of additional capacity at stage 3. The total expected cost of the flexible design is calculated using Eq. (15). Table IV shows the optimal design for the HRES with design with capacity expansion in 2027 and 2032. Similar to case 1, the initial design of the system will satisfy demand for 2017-2026 considering all of the 10,000 simulations for demand. The initial optimal design is the same as in case 1. In 2027, the capacity of the wind turbines and battery are expanded for the high and medium demand profiles, but not as much as in case 1. In case 2, the total capacity in 2027 should fulfill demand from 2027-2031, but in case 1, it is necessary to expand capacity to meet demand from 2027-2036. If demand is low in stage 2, none of the system components' capacities will be expanded, which is similar to the case 1. Instead, the needed electricity will be purchased from the market with diesel.

Given the initial design in stage 1 and the decisions about expanding capacity in stage 2, the additional capacity is optimized for high, medium, and low demand profiles for 2032-2036. The capacity expansion amounts at stage 2 are the same for the medium and high demand profiles in 2027-2031 (see Table IV), and the planning for stage 3 given medium demand in stage 2 is the same as the planning for stage 3 given high demand in stage 2. Additional capacity for the battery should be added for high demand profiles in stage 3. If demand is high or medium in stage 2 and low in stage 3, the capacity of the hydrogen tank should be expanded. In these scenarios, the total amount of energy generated by the wind turbine and solar panel exceeds demand, and the surplus amount should be reserved in the hydrogen tank.

There should be no expansion in stage 2 if demand is low. If demand remains low in stage 3, there should not be any expansion either. If demand is high or medium in stage 3, the decision maker should add additional capacity to the wind turbine and the battery. The capacity increases for high and medium demand in stage 3 are very similar.

Bayesian optimization algorithm and the simulation model are implemented in MATLAB software. The time it takes to run Bayesian Optimization algorithm for the static problem (design without capacity expansion) is approximately 30 minutes; However, it takes approximately 210 minutes to find optimal solution for the two stage capacity expansion problem.

3) Value of flexibility or capacity expansion: In the first case of designing with capacity expansion, less capacity is needed for the first 10 years operation because demand does not increase significantly. If demand is high or medium in month 120, the designer should choose to increase the capacity of the wind turbine and battery. In case 2, the initial capacity should be similarly expanded for the wind turbine and battery if demand is high or

TABLE III: The optimal design of the HRES with capacity expansion for the 1 stage case

Component	Initial design (Giga watt)	Stage 2		
		High demand	Medium Demand	Low Demand
Solar panel	263	0	0	0
Wind turbine	31	128	98	0
Battery	17	54	39	0
Electrolyzer	230	0	0	0
Hydrogen tank	616	0	0	0
Fuel cell	68	0	0	0
Expected Cost ($ trillion)	20.55	12.22	7.63	7.08
Total Cost ($ trillion)	27.22			
Value of capacity expansion ($ trillion)	13.44			

TABLE IV: The optimal design of the HRES with capacity expansion in design for the two stages case

Component	Initial design	Stage 2			Stage 3 given high or medium demand in stage 2			Stage 3 given low demand in stage 2		
		High demand	Medium demand	Low demand	High demand	Medium demand	Low demand	High demand	Medium demand	Low demand
Solar panel	263	0	0	0	0	0	0	0	0	0
Wind turbine	31	62	62	0	0	0	0	106	107	0
Battery	17	25	25	0	50	0	0	47	56	0
Electrolyzer	230	0	0	0	0	0	0	0	0	0
Hydrogen tank	616	0	0	0	0	0	211	0	0	0
Fuel cell	68	0	0	0	0	0	0	0	0	0
Cost ($ trillion)	20.55	6.11	3.33	3.19	6.38	4.72	0.41	7.5	4.3	2.91
Total Cost ($ trillion)	26.52									
Value of capacity expansion($ trillion)	14.14									

medium in stage 2. Differences occur between the two cases because case 2 has an additional stage to plan for in years 2032-2036. If demand is high or medium in stage 2 and low in stage 3, the capacity of the hydrogen tank should be expanded to have more space for the excess amount of energy generated. If demand is low in stage 2 and high or medium in stage 3, the capacity of the battery and wind turbine should be expanded to meet the rising demand because no expansion occurred in stage 2.

The expected discounted cost of design with capacity expansion for case 2 is $26.52 trillion which is less than the expected cost in case 1. In case 2, the designers have two options to exercise, one in 2027 and one in 2032. Delaying a decision on expanding capacity to 2032 allows the designers to take advantage of the time value of money.

Figure 5 shows that the expected cost decreases as more design modifications are included. In reality, there may be an increase in the initial investment cost that will allow the designers to easily expand capacity in the future. If that increase in the initial investment cost to enable the possible design modification in the future is less than $13.44 trillion for the one-stage modification or $14.14 trillion for the two-stage modification, then the designer should spend the money to have that option available to him or her in the future. These two values show the value of adding flexibility or capacity expansion to the design of HRES.

Flexibility usually requires an upfront cost in order to be able to pursue the flexible alternatives. However, our methodology calculates the expected cost of design with flexibility (or capacity expansion). The difference between the expected costs of the design with capacity expansion and without capacity expansion is the maximum amount that the designer should pay to have a flexible option. This article only studies one type of flexibility modeling (i.e., capacity expansion) for the HRES. Future research can study other flexibility modeling ideas such as abandoning the system permanently or switching design configurations [27].

IV. Conclusion

This paper has presented a method to incorporate the demand uncertainty into the multi-stage design decisions of an HRES. The HRES is composed of six components: solar panel, wind turbine, battery, electrolyzer, hydrogen tank, and fuel cell. The electricity demand data for California for over a 20-year period is simulated with an ARIMA time series model. The Bayesian optimization algorithm identifies the optimal design of the HRES by minimizing the expected discounted cost considering the demand for electricity for California and other uncertain cost parameters over 20 years. A design with capacity expansion is conducted in two cases: a single design modification and two opportunities to modify the design.

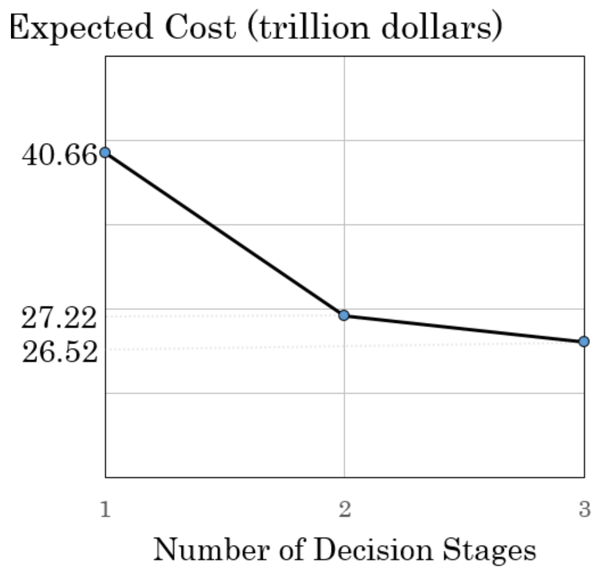


Fig. 5: The expected cost vs. decision stages of the HRES

The results show that a single design modification 10 years after system deployment reduces the system's expected discounted cost by 33%. Including a second design modification would reduce the expected cost by an additional 3%.

This paper makes an important contribution to the literature of flexible design by measuring the value of capacity expansion in a complex engineered systems which require the use of computationally expensive simulations to evaluate the objective function. The model optimizes the design of engineered system by using probability distributions to forecast highly uncertain demand 20 years into the future. A multi-stage decision-making model algorithm is developed to evaluate the value of flexibility or capacity expansion. The algorithm is tested considering three categories at each stage. Future studies could further analyze and find the optimal number of divisions at each stage.

Future research could also include the time to exercise an option as a decision variable. For example, Kucuksayacigil and Min [57] use real options analysis to find the optimal time to enlarge a ship after the ship is designed. If simulation optimization is required to optimize over many periods, such as 240 months, it is time consuming and even impossible with today's CPUs to optimize the mathematical model 240 times rolling back from the end for the 10,000 simulations. A heuristic model could be developed to find the optimal time to review the initial design of HRES.

The proposed algorithm for multi-stage decision-making design problems can be applied to any complex engineered system such as jet engines design and self-driving cars. These complex engineered systems require optimizing high-dimensional, computationally expensive objective functions in a highly uncertain environment. Our proposed algorithm can potentially help designers to design complex engineered systems flexible to future uncertain situations.

Acknowledgment

References

[1] R. De Neufville and S. Scholtes, Flexibility in engineering design. MIT Press, 2011.
[2] N. Kang, A. E. Bayrak, and P. Y. Papalambros, "A real options approach to hybrid electric vehicle architecture design for flexibility," in ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, 2016, pp. V02AT03A041–V02AT03A041.
[3] X. Zhang, H.-Z. Huang, Z. Wang, Y. Liu, and H.-W. Xu, "A new method for achieving flexibility in hierarchical multilevel system design," Concurrent Engineering, vol. 19, no. 2, pp. 187–196, 2011.
[4] M.-A. Cardin, Q. Xie, T. S. Ng, S. Wang, and J. Hu, "An approach for analyzing and managing flexibility in engineering systems design based on decision rules and multistage stochastic programming," IISE Transactions, vol. 49, no. 1, pp. 1–12, 2017.
[5] J. H. Saleh, D. E. Hastings, and D. J. Newman, "Flexibility in system design and implications for aerospace systems," Acta astronautica, vol. 53, no. 12, pp. 927–944, 2003.
[6] P. N. Koch, T. W. Simpson, J. K. Allen, and F. Mistree, "Statistical approximations for multidisciplinary design optimization: the problem of size," Journal of aircraft, vol. 36, no. 1, pp. 275–286, 1999.
[7] L. Gu, "A comparison of polynomial based regression models in vehicle safety analysis," in ASME Design Engineering Technical Conferences, ASME Paper No.: DETC/DAC-21083, 2001.
[8] M.-A. Cardin, "Enabling flexibility in engineering systems: a taxonomy of procedures and a design framework," Journal of Mechanical Design, vol. 136, no. 1, p. 011005, 2014.
[9] S. D. Wu, M. Erkoc, and S. Karabuk, "Managing capacity in the high-tech industry: A review of literature," The engineering economist, vol. 50, no. 2, pp. 125–158, 2005.
[10] R. L. Smith, "Optimal expansion policies for the deterministic capacity problem," The Engineering Economist, vol. 25, no. 3, pp. 149–160, 1979.
[11] G. C. Prueitt and C. S. Park, "Phased capacity expansion-using continuous distributions to model prior beliefs," The Engineering Economist, vol. 42, no. 2, pp. 91–110, 1997.
[12] G. C. Philip and J. M. Liittschwager, "An optimal capacity expansion model with economies of scale," The Engineering Economist, vol. 24, no. 4, pp. 195–216, 1979.
[13] E. E. Karsak and C. O. ÖZOGUL, "An options approach to valuing expansion flexibility in flexible manufacturing system investments," The Engineering Economist, vol. 47, no. 2, pp. 169–193, 2002.
[14] J. C. Berretta and F. Mobasheri, "An optimal strategy for capacity expansion," The Engineering Economist, vol. 17, no. 2, pp. 79–98, 1972.
[15] J. C. Ammons and L. McGinnis, "A generation expansion planning model for electric utilities," The Engineering Economist, vol. 30, no. 3, pp. 205–226, 1985.
[16] J. Wang, M. Shahidehpour, Z. Li, and A. Botterud, "Strategic generation capacity expansion planning with incomplete information," IEEE Transactions on Power Systems, vol. 24, no. 2, pp. 1002–1010, 2009.
[17] A. Khodaei, M. Shahidehpour, and S. Kamalinia, "Transmission switching in expansion planning," IEEE Transactions on Power Systems, vol. 25, no. 3, pp. 1722–1733, 2010.
[18] E. Hajipour, M. Bozorg, and M. Fotuhi-Firuzabad, "Stochastic capacity expansion planning of remote microgrids with wind farms and energy storage," IEEE transactions on sustainable energy, vol. 6, no. 2, pp. 491–498, 2015.
[19] E. Gil, I. Aravena, and R. Cárdenas, "Generation capacity expansion planning under hydro uncertainty using stochastic mixed integer programming and scenario reduction," IEEE Transactions on Power Systems, vol. 30, no. 4, pp. 1838–1847, 2014.
[20] J. Hu, P. Guo, and K. L. Poh, "Flexible capacity planning for engineering systems based on decision rules and differential evolution," Computers & Industrial Engineering, vol. 123, pp. 254–262, 2018.

[21] J. Hu and M.-A. Cardin, “Generating flexibility in the design of engineering systems to enable better sustainability and lifecycle performance,” Research in Engineering Design, vol. 26, no. 2, pp. 121–143, 2015.

[22] R. D. MacKinnon, “The concept of flexibility and urban systems.” 1968.

[23] M. A. Fox and B. P. Yeh, “Intelligent kinetic systems in architecture,” in Managing interactions in smart environments. Springer, 2000, pp. 91–103.

[24] M. Amram, N. Kulatilaka et al., “Real options:: Managing strategic investment in an uncertain world,” OUP Catalogue, 1998.

[25] A. Raouf and M. Ben-Daya, Flexible manufacturing systems: recent developments. Elsevier, 1995, vol. 23.

[26] J. R. Highsmith, Adaptive software development: a collaborative approach to managing complex systems. Addison-Wesley, 2013.

[27] M.-A. Cardin, R. de Neufville, and D. M. Geltner, “Design catalogs: a systematic approach to design and value flexibility in engineering systems,” Systems engineering, vol. 18, no. 5, pp. 453–471, 2015.

[28] M. Deshmukh and S. Deshmukh, “Modeling of hybrid renewable energy systems,” Renewable and Sustainable Energy Reviews, vol. 12, no. 1, pp. 235–249, 2008.

[29] A. K. Kaviani, G. Riahy, and S. M. Kouhsari, “Optimal design of a reliable hydrogen-based stand-alone wind/pv generating system, considering component outages,” Renewable energy, vol. 34, no. 11, pp. 2380–2390, 2009.

[30] M. Sharafi and T. Y. ELMekkawy, “Multi-objective optimal design of hybrid renewable energy systems using pso-simulation based approach,” Renewable Energy, vol. 68, pp. 67–79, 2014.

[31] A. Khalilnejad, A. Sundararajan, and A. I. Sarwat, “Optimal design of hybrid wind/photovoltaic electrolyzer for maximum hydrogen production using imperialist competitive algorithm,” Journal of Modern Power Systems and Clean Energy, vol. 6, no. 1, pp. 40–49, 2018.

[32] S. S. Sadati, E. Jahani, and O. Taylan, “Technical and economic analyses for sizing pv power plant with storage system for metu ncc,” in ASME 2015 International Mechanical Engineering Congress and Exposition. American Society of Mechanical Engineers, 2015, pp. V06BT07A044–V06BT07A044.

[33] R. Dufo-López and J. L. Bernal-Agustín, “Multi-objective design of pv–wind–diesel–hydrogen–battery systems,” Renewable energy, vol. 33, no. 12, pp. 2559–2572, 2008.

[34] Y. Chen, M. W. Hoffman, S. G. Colmenarejo, M. Denil, T. P. Lillicrap, and N. de Freitas, “Learning to learn for global optimization of black box functions,” 2018.

[35] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas, “Taking the human out of the loop: A review of bayesian optimization,” Proceedings of the IEEE, vol. 104, no. 1, pp. 148–175, 2016.

[36] R. Martinez-Cantin, “Bayesopt: A bayesian optimization library for nonlinear optimization, experimental design and bandits,” The Journal of Machine Learning Research, vol. 15, no. 1, pp. 3735–3739, 2014.

[37] D. J. Lizotte, Practical bayesian optimization. University of Alberta, 2008.

[38] J. Snoek, H. Larochelle, and R. P. Adams, “Practical bayesian optimization of machine learning algorithms,” in Advances in neural information processing systems, 2012, pp. 2951–2959.

[39] J. Mockus, V. Tiesis, and A. Zilinskas, “Toward global optimization, volume 2, chapter bayesian methods for seeking the extremum,” 1978.

[40] Z. Wang, M. Zoghi, F. Hutter, D. Matheson, N. De Freitas et al., “Bayesian optimization in high dimensions via random embeddings.” in IJCAI, 2013, pp. 1778–1784.

[41] “arima class, https://www.mathworks.com/help/econ/arima-class.html, matlab central file exchange.”

[42] J. W. Taylor, “Short-term electricity demand forecasting using double seasonal exponential smoothing,” Journal of the Operational Research Society, vol. 54, no. 8, pp. 799–805, 2003.

[43] D. Akay and M. Atak, “Grey prediction with rolling mechanism for electricity demand forecasting of turkey,” Energy, vol. 32, no. 9, pp. 1670–1675, 2007.

[44] S. S. Sadati, E. Jahani, O. Taylan, and D. K. Baker, “Sizing of photovoltaic-wind-battery hybrid system for a mediterranean island community based on estimated and measured meteorological data,” Journal of Solar Energy Engineering, vol. 140, no. 1, p. 011006, 2018.

[45] A. Arima, “A. arima and f. iachello, ann. phys.(ny) 99, 253 (1976),” Ann. Phys.(NY), vol. 99, p. 253, 1976.

[46] S. Yuan, A. S. Kocaman, and V. Modi, “Benefits of forecasting and energy storage in isolated grids with large wind penetration–the case of sao vicente,” Renewable energy, vol. 105, pp. 167–174, 2017.

[47] M. De Felice, A. Alessandri, and P. M. Ruti, “Electricity demand forecasting over italy: Potential benefits using numerical weather prediction models,” Electric Power Systems Research, vol. 104, pp. 71–79, 2013.

[48] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, “Day-ahead electricity price forecasting using the wavelet transform and arima models,” IEEE transactions on power systems, vol. 20, no. 2, pp. 1035–1042, 2005.

[49] T. Jakaša, I. Andročec, and P. Sprčić, “Electricity price forecasting—arima model approach,” in 2011 8th International Conference on the European Energy Market (EEM). IEEE, 2011, pp. 222–225.

[50] M. Zhou, Z. Yan, Y. Ni, G. Li, and Y. Nie, “Electricity price forecasting with confidence-interval estimation through an extended arima approach,” IEE Proceedings-Generation, Transmission and Distribution, vol. 153, no. 2, pp. 187–195, 2006.

[51] Z. Tan, J. Zhang, J. Wang, and J. Xu, “Day-ahead electricity price forecasting using wavelet transform combined with arima and garch models,” Applied Energy, vol. 87, no. 11, pp. 3606–3610, 2010.

[52] M. Zhou, Z. Yan, Y. Ni, and G. Li, “An arima approach to forecasting electricity price with accuracy improvement by predicted errors,” in IEEE Power Engineering Society General Meeting, 2004. IEEE, 2004, pp. 233–238.

[53] J. Che and J. Wang, “Short-term electricity prices forecasting based on support vector regression and auto-regressive integrated moving average modeling,” Energy Conversion and Management, vol. 51, no. 10, pp. 1911–1917, 2010.

[54] U. E. I. Administration, “https://www.eia.gov/opendata/qb.php?category=38.”

[55] C. Yang, S. Yeh, S. Zakerinia, K. Ramea, and D. McCollum, “Achieving california’s 80% greenhouse gas reduction target in 2050: Technology, policy and scenario analysis using ca-times energy economic systems model,” Energy Policy, vol. 77, pp. 118–130, 2015.

[56] S. W. Miller, T. W. Simpson, M. A. Yukish, G. Stump, B. L. Mesmer, E. B. Tibor, C. L. Bloebaum, and E. H. Winer, “Toward a value-driven design approach for complex engineered systems using trade space exploration tools,” in ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, 2014, pp. V02AT03A052–V02AT03A052.

[57] F. Kucuksayacigil and K. J. Min, “Value of jumboization in ship design: A real options approach,” in IIE Annual Conference. Proceedings. Institute of Industrial and Systems Engineers (IISE), 2017, pp. 1595–1600.