# Parallel Computing for Power System Climate Resiliency: Solving a Large-Scale Stochastic Capacity Expansion Problem with mpi-sppy

Tomas Valencia Zuluaga,
Shmuel S. Oren
Industrial Engineering & Operations Research
University of California, Berkeley
Berkeley, CA, USA
{tvalenciaz, shmuel}@berkeley.edu

Amelia Musselman,
Jean-Paul Watson
Cyber and Infrastructure Resilience
Lawrence Livermore National Laboratory
Livermore, CA, USA
{musselman5, watson61}@llnl.gov

***Abstract*—We propose a nodal stochastic generation and transmission expansion planning model that incorporates the output from high-resolution global climate models through load and generation availability scenarios. We implement our model in Pyomo and perform computational studies on a realistically-sized test case of the California electric grid in a high performance computing environment. We propose model reformulations and algorithm tuning to efficiently solve this large problem using a variant of the Progressive Hedging Algorithm. We utilize the parallelization capabilities and overall versatility of *mpi-sppy*, exploiting its hub-and-spoke architecture to concurrently obtain inner and outer bounds on an optimal expansion plan. Initial results show that instances with 360 representative days on a system with over 8,000 buses can be solved to within 5% of optimality in under 4 hours of wall clock time, a first step towards solving a large-scale power system expansion planning problem across a wide range of climate-informed operational scenarios.**

***Index Terms*—Stochastic programming, capacity expansion, parallel computing, climate resiliency**

## I. Introduction

Widespread penetration of renewable, intermittent, and decentralized generation resources is rapidly transforming the power grid and increasing its sensitivity to weather. Maintaining a resilient power grid as this transition unfolds will require strategic infrastructure investments. The recent *Net-Zero America* study [1] estimates that achieving carbon neutrality by 2050 will require quadrupling previous average annual build rates for renewable generation and increasing the total installed transmission capacity by a factor between 3 and 5. Conventional modeling tools for generation and transmission expansion planning, based on historical data, are inadequate to inform investment decisions under this changing environment. Incorporating climate projections into expansion planning tools can improve their relevance to decision makers, but the results are only meaningful if the uncertainty associated with these projections is also considered. Stochastic optimization is well-suited for this purpose, but often leads to computationally challenging optimization problems.

Developing actionable expansion plans for an efficient and resilient power system requires both sufficiently high geographical resolution to accurately represent power system components and a sufficiently large representative set of scenarios to encompass all relevant potential weather impacts. The curse of dimensionality usually forces planners to face a tradeoff between these two dimensions.

The first computational challenge arises due to the size of the power system model. It is common to use very coarse geographic resolution models, representing entire states or nations as a single node in the power network, for theoretical power system planning studies. Obtaining actionable decisions from these unrealistic representations of the power grid requires iteratively solving increasingly high-resolution optimization models for each subregion of the grid. Solving power system planning models in this way not only leads to sub-optimal solutions but also underestimates the variability of wind and solar resources and load by aggregating these resources across a large area. To address the need for a realistic power system model suitable for research studies, the California Test System (CATS) was recently developed [2]. This system, which we adopt and extend for the tests presented in this paper, includes over 8,000 buses for the state of California. For comparison, a model often used for expansion plans in the literature has 240 buses for the entire U.S. Western Interconnection [3]. Instances of capacity expansion models that maintain the full detail of the transmission network, and thus have comparable sizes, exist in the literature, but are solved via deterministic scenario planning [4], or, if considering uncertainty, for a significantly

Submitted to the 23rd Power Systems Computation Conference (PSCC 2024). This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 and was supported by the LLNL-LDRD Program under Project 22-SI-008, and by the Advanced Grid Modeling Program of the Office of Electricity of the U.S. Department of Energy. The authors would like to thank their LLNL colleagues Minda Monteagudo and Matthew Signorotti for their invaluable contributions in translating the climate data projections into relevant power system input parameters. The authors would also like to thank Gurobi for providing the academic license used to run all tests.

reduced number of scenarios [5].

The second computational challenge is in adequately capturing the uncertainty, which we address using stochastic optimization. Even without consideration of climate-dependent uncertainties, stochastic optimization has been shown to have significant economic benefits for power grid capacity expansion planning when compared to deterministic optimization and heuristic scenario planning [6], [7]. Climate uncertainty underscores these benefits. A rapidly changing climate requires climate model projections, rather than just historical data, to be fed into capacity expansion models. However, despite recent advances in global circulation model downscaling algorithms to obtain multi-decadal, high-resolution weather projections, the uncertainty associated with these projections is substantial, and so can be the potentially adverse effect of picking a single deterministic timeseries to inform investment plans.

Despite this necessity, the usage of stochastic optimization in actual planning processes is still modest, with computational limitations being an often-cited reason for preferring deterministic models. Muñoz et al. [6], [8] and Go et al. [7] lay the ground-work for using stochastic programming for capacity expansion planning, addressing the aforementioned tradeoff through model simplifications. In [8], a scalable stochastic expansion model is proposed, using decomposition by scenarios, with tests conducted on a universe of 8,760 one-hour scenarios that is reduced to 100-500 scenarios after clustering, on the geographically simplified 240-bus WECC test case. However, interconnection between time periods are neglected, which does not allow storage or other inter-temporal resources to be represented. Inter-temporal storage decisions are included in [7], but only a small, 24-bus, test-case with 5 scenarios is solved.

In the examples above, the scenarios of the stochastic optimization problem are obtained by constructing a probability space from historical data, then reducing that probability space to a sample space of representative days or hours through statistical clustering techniques. A similar approach can be taken by substituting future weather projections for historical data, which is what we do in this work. We propose thus a climate-dependent stochastic generation, transmission, and storage capacity expansion model adapted from the model proposed in [7] and solve this model for a realistically-sized test case of California, leveraging *mpi-sppy* and high performance computing (HPC) resources at Lawrence Livermore National Laboratory.

*mpi-sppy* [9] is a recently developed extension of the *Pyomo* [10], [11] optimization package in Python, specifically designed to solve stochastic optimization problems in HPC environments using the Message Passing Interface (MPI). Besides providing a simple, parallel, way of implementing decomposition algorithms to solve stochastic programs using the Python programming language, the *hub-and-spoke* architecture of mpi-sppy addresses two known shortcomings of the Progressive Hedging Algorithm in the mixed-integer context: the lack of optimality bounds during the execution of the algorithm and the related lack of a good termination criterion.

We solve our model of the aforementioned high spatial-resolution, realistic test system of the California electricity grid [2], with a large number of scenarios derived from the California regional refinement model (CARRM) of the Energy Exascale Earth System Model (E3SM) using a 3km-resolution grid [12]. In Section II, we summarize the capacity expansion model proposed. In Section III, the key features of the solution approach implemented are described. Finally, results are presented and discussed in Section IV.

## II. Model description

Our capacity expansion model is an adaptation of the model proposed in [7]. Due to space limitations, we provide a brief general description of the model with details given only for features that differ from the reference model. We refer the interested reader to [7] for the full model description.

### A. General description

Our capacity expansion model is a two-stage stochastic program that co-optimizes generation, transmission and storage investments in the first stage and solves a multi-period optimal power flow (OPF) problem in the second stage. Second-stage suproblems, i.e. representative days, differ from each other in the hourly system demand and generation availability.

Renewable generation and storage investments are modeled as continuous (installed capacity), thermal generation investments as integer (number of installed units), and transmission investments as binary (build or no-build for candidate lines) variables. All investment costs are assumed to be linear. Storage investments consist of two decisions: energy storage capacity (in MWh, which can be interpreted as storage duration) and instantaneous power charging/discharging capacity (in MW). These two decisions are made independently.

At this stage of this work, the second-stage OPF is modeled with a transportation relaxation of power flow incorporating transmission losses. Forthcoming extensions shall use the more common linearized DC-OPF. Ramping and start-up constraints are disregarded. Storage levels are assumed to be cyclic (level at the end of the last period must equal the initial storage level). The constraint to impede simultaneous charging and discharging of storage units is omitted because this was found in [7] to have minimal solution impact while increasing computational complexity.

In the following subsections, we describe in more detail the set of investment candidates, our model for transmission losses, and the renewable portfolio standard (RPS) constraints, all of which deviate from the model in [7].

### B. Investment candidates

In the model description of [7], the set of generation and storage investment candidates consists of all combinations of candidate technologies and power system buses. This set is later reduced to an explicit list of candidate units when the problem data is specified. We make this methodology precise by considering all combinations of technologies and buses subject to three *maximum potential capacity* constraints. The

three constraints limit the total installed capacity for each storage/generation type at each bus, for each type across all buses, and for each bus across all types. In actual instances of the model, the maximum potential capacity is zero for many generator type, location combinations, allowing the number of variables included in the instantiation of the model to remain tractable.

### C. Transmission losses

As is common when performing planning studies, the model in [7] disregards transmission losses. However, ignoring these losses can significantly affect the solution obtained from the model, especially for systems with renewable generation, in which power may need to be transported over very large distances [13], [14]. Accurately capturing the quadratic nature of transmission losses without losing the model's linearity requires adding several linear segments for each transmission line, which greatly increases the number of variables in the model. To avoid this increase, we consider a single-segment linear approximation. Specifically, we consider each transmission branch $\ell$ to have a constant efficiency $\eta_\ell^{\mathcal{L}}$ given by

$$\eta_\ell^{\mathcal{L}} = 1 - P_\ell^{\mathcal{L}} r_\ell \left(1 + \frac{r_\ell^2}{x_\ell^2}\right), \tag{1}$$

where $x_\ell$ is the reactance, $r_\ell$ is the resistance, and $P_\ell^{\mathcal{L}}$ is the long-term branch rating of branch $\ell$, all in p.u. This linear approximation, based on Fitiwi et al. [14], will overestimate losses in transmission branches that are loaded below their capacity, leading to somewhat conservative solutions. In contrast, assuming lossless transmission across long distances will significantly underestimate transmission capacity needs.

In order to correctly assign losses to buses, we also deviate from [7] by using two non-negative variables $p_\ell^{\mathcal{L}+}, p_\ell^{\mathcal{L}-}$ to represent branch flows, as illustrated in Fig. 1. Note that formulation (b) only represents formulation (a) accurately if $p_\ell^{\mathcal{L}+} \cdot p_\ell^{\mathcal{L}-} = 0$. Otherwise, model (b) will have higher losses than the corresponding model (a) with same net outgoing branch flow.

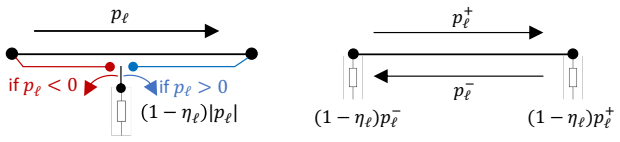


Fig. 1. Proposed representation of flow in transmission branches. In model (a), depending on the direction of flow, the switch determines the end of the branch to which losses are assigned. In model (b), the same is achieved by breaking the sign-unconstrained flow $p_\ell$ into positive and negative parts, representing flow in each direction, so that $p_\ell^{\mathcal{L}} = p_\ell^{\mathcal{L}+} - p_\ell^{\mathcal{L}-}$. Model (b) only represents model (a) accurately if $p_\ell^{\mathcal{L}+} \cdot p_\ell^{\mathcal{L}-} = 0$.

In other words, making $p_\ell^{\mathcal{L}+}$ and $p_\ell^{\mathcal{L}-}$ simultaneously nonzero artificially increases the losses in a transmission element, which we call spurious losses. This phenomenon can be prevented, at high computational cost, by introducing a binary variable for each branch and time period to enforce $p_\ell^{\mathcal{L}+} \cdot p_\ell^{\mathcal{L}-} = 0$. This is normally considered unnecessary, because spurious losses usually increase costs and would hence be naturally avoided by the optimization model. However, during certain periods, if the locational marginal prices (LMP) at the ends of a transmission element are negative, creating spurious losses will be optimal.

In the absence of policy incentives, negative LMPs are related with congestion during unusual loading conditions and are rare [14]. In [13], they have been found to occur in less than 0.2% of operating hours. The introduction of RPS constraints can create incentive for artificial load loss (through transmission and storage loops) if not implemented properly. A way of addressing this issue without increasing the computational burden is discussed in the following section.

### D. Renewable Portfolio Standards

RPS constraints represent policies in place in certain jurisdictions, whereby it is mandated that at least some fraction $\rho^{RPS}$ of annual energy produced must be generated with qualified renewable resources. Scenarios correspond to representative days of the target year, not to different realizations thereof. Therefore, the RPS constraint cannot be enforced for each scenario. Rather, it is formulated using the expectation across scenarios as follows:

$$\sum_{\omega\in\Omega} \rho_\omega \sum_{\substack{b\in\mathcal{B}, g\in\mathcal{G}^R \\ t\in\mathcal{T}}} p_{b,g,t,\omega}^{\mathcal{G}} \geq \rho^{\text{RPS}} \sum_{\omega\in\Omega} \rho_\omega \sum_{\substack{b\in\mathcal{B}, g\in\mathcal{G} \\ t\in\mathcal{T}}} p_{b,g,t,\omega}^{\mathcal{G}}, \tag{2}$$

where $\omega$ indexes the discrete set of scenarios $\Omega$ with probabilities $\rho_\omega$, $b$ the set of buses $\mathcal{B}$, $g$ the set of generator types $\mathcal{G}$, and its subset of renewable generator types $\mathcal{G}^R$ and $t$ indexes the set of time periods $\mathcal{T}$ in the second-stage subproblem. $p_{b,g,t,\omega}^{\mathcal{G}}$ is the generation output of generator type $g$ at bus $b$ during period $t$ in scenario $\omega$. Constraint (2) presents a challenge for both computational tractability and model validity.

First, the expectation constraint couples the second-stage variables across scenarios sucth that it is no longer truly a stochastic program. This is not a computational issue in [7] because the extensive form of the problem is solved. However, when implementing decomposition algorithms as proposed here, second-stage decision variables must be independent across scenarios. A possible solution is to dualize the constraint and add the corresponding Lagrangian term with multiplier $\lambda^{RPS}$ to the objective, to create a reformulation that is again decomposable by scenario. This approach is justified by the existence of a Renewable Energy Certificates (REC) market where utilities may trade surplus and shortfall of renewable energy production across days to satisfy the RPS mandate. Nonetheless, it poses the additional problem of finding the right value for $\lambda^{RPS}$, the REC price. In [8], where this approach is taken, three methods are proposed: (a) using the value from real-world REC markets, (b) using the value of the dual variable of (2) in the solution to the linear relaxation of the stochastic program in extensive form (impractical for

very large instances like ours), and (c) performing a sensitivity analysis until the achieved proportion of renewable energy in total production is satisfactory. We implement the third option.

The second challenge, to which we alluded in the previous section, is inducing spurious losses through negative LMPs. During days with low demand and high renewable availability, the REC market can depress LMPs and make them negative in certain parts of the grid. If this is widespread, the accuracy of the model is compromised, since there is an incentive to under-invest in renewable generation facilities, and meet RPS targets by way of creating artificial loads during such days instead. Tests in our model showed that if left uncontrolled, spurious losses could increase to over 50% of total demand during certain operating hours. Such an unrealistic behavior would put the validity of the whole model into question.

To address this issue while avoiding the computational burden of adding binaries to prevent spurious losses, we make the RPS constraint a soft constraint enforced at each scenario, as follows:

$$p^{NC}_{\omega} + \sum_{\substack{b\in\mathcal{B}, g\in\mathcal{G}^R \\ t\in\mathcal{T}}} p^{\mathcal{G}}_{b,g,t,\omega} \geq \rho^{\text{RPS}} \sum_{\substack{b\in\mathcal{B}, g\in\mathcal{G} \\ t\in\mathcal{T}}} p^{\mathcal{G}}_{b,g,t,\omega} \;\forall\, \omega \in \Omega \quad (3)$$

$p^{NC}_{\omega} \geq 0$ is the scenario's shortfall when there is one, i.e. the RPS target non-compliance of each scenario. A term $\lambda^{RPS} p^{NC}_{\omega}$ is added to each subproblem's objective. This model corresponds to a case where there is no REC market; instead, penalties are assessed at the end of each day and charged to utilities found to be out of compliance. Our model is likely to over-incentivize renewable investments and over-estimate expected operational costs, compared to an implementation of the RPS policy as an expectation constraint, but is a first step towards a formulation that can allow transmission losses while tractably handling policy-induced spurious losses.

## III. Solution approach

As mentioned previously, we leverage the implementation of the Progressive Hedging Algorithm (PHA) in mpi-sppy, and make adaptations as needed to solve the capacity expansion problem posed. Detailed descriptions of PHA can be found in [15] and [16], among many other references. In this section, we highlight the main features of our implementation that differ from conventional implementations of PHA.

### A. Accelerating subproblems

In PHA, the stochastic program is decomposed by scenario. As for any decomposition algorithm, the key to a successful implementation is that each subproblem can be solved quickly. This section highlights measures undertaken to ensure that individual subproblems are quick to solve.

**Simplifying transmission losses** The model described in section II-C adds one variable to the second-stage subproblem for each transmission branch and time period (approximately 260,000 variables in our test case), which can have a large impact on subproblem solution time. Furthermore, note that for branches with negligible losses, i.e. $\eta^{\mathcal{L}}_{\ell} \approx 1$, replacing $p^{\mathcal{L}}_{\ell}$ with $p^{\mathcal{L}+}_{\ell}, p^{\mathcal{L}-}_{\ell}$ increases degeneracy[1], which can be problematic.

To mitigate these undesired effects, we propose partitioning the set of transmission branches into *lossy* and *lossless* branches. Lossy branches are modeled as described above, while flow through lossless branches is modeled with a single sign-unconstrained variable $p^{\mathcal{L}}_{\ell}$. Testing across different loading configurations in our test case, we have found that approximating all branches with $1 - \eta^{\mathcal{L}}_{\ell} < 5 \cdot 10^{-3}$ as lossless branches reduces the number of additional variables by over 50% and the subproblem solution time by nearly 50%, while still accounting for more than 80% of the losses in the system.

**Linearize quadratic term** Despite recent advances in commercial solvers, quadratic mixed integer programs remain harder to solve than mixed integer linear programs (MILPs) and the quadratic (also called *proximal*) terms introduced by PHA can considerably delay the subproblem solution times. To reduce this effect, we linearize the proximal terms in the PHA subproblems. For binary variables, the quadratic term automatically reduces to a linear term, so no special linearization is necessary. For integer and continuous variables, an extra variable is added to represent the squared variables, and the quadratic penalty is under-approximated by successive linear cuts added as needed during the execution of the algorithm. This linearization procedure is already included in the current mpi-sppy distribution.

**Accelerating crossover** The branch and cut algorithm implemented by commercial MILP solvers like Gurobi requires a basic feasible solution to the linear relaxation of the problem – the so-called root relaxation on which the algorithm branches. By design, the Simplex Algorithm returns such a solution. For problems of large size, like the one considered here, interior point algorithms are much faster, but require a crossover phase to convert their solution into an extreme point optimal solution. The crossover process can be accelerated by reducing degeneracy in the problem formulation. In our case, a few simple model improvements like adding a reference bus and simplifying the model for branches with negligible losses as described above were used to improved crossover time.

**Cleaning the input data** Real (or realistic) power system data can lead to ill-conditioned matrices and numerically unstable models. An undesired byproduct of the geographically accurate test system of [2] is a large number of very short transmission branches, which have positive, but very small reactance and resistance values and reinforcement costs. These do not significantly affect the optimal solution or cost, but can become a numerical nuisance for the optimization problem. To reduce this effect, certain buses in the network were collapsed and their incident branches absorbed into their neighbors. Although this would seem to go against the stated goal of increasing geographic resolution to adequately assess climate impacts, it was only performed when doing so did

[1] Observe that for given $\hat{p}^{\mathcal{L}+}_{\ell}, \hat{p}^{\mathcal{L}-}_{\ell}$, any $p^{\mathcal{L}+}_{\ell}, p^{\mathcal{L}-}_{\ell}$ such that $p^{\mathcal{L}+}_{\ell} - p^{\mathcal{L}-}_{\ell} = \hat{p}^{\mathcal{L}+}_{\ell} - \hat{p}^{\mathcal{L}-}_{\ell}$ and such that $p^{\mathcal{L}+}_{\ell} + p^{\mathcal{L}-}_{\ell}$ remains below the maximum branch capacity will be a feasible solution with same objective value.

not affect the resulting power flow. This has already improved the subproblem solution time for our current model, but is expected to have a higher impact on the DC-OPF formulation.

### B. Progressive Hedging settings

**Setting $\rho$** The value chosen for the parameter $\rho$, the coefficient associated with the quadratic penalty term and the weight update step, is known to have a great impact on the convergence of the PHA. Although the standard description of the PHA assumes that the same parameter $\rho$ is used for all non-anticipative variables, it has been found that better results can be obtained if the parameter is made variable-dependent. In [8], generator investment variables are assigned different values of $\rho$ from transmission investment variables, so as to maintain an approximately constant ratio of penalty/(investment $). In [17], it is found that even better results can be obtained if the value of $\rho$ is determined individually for each variable after assessing the level of agreement in the optimal solution across scenarios at the first PHA iteration. The authors call this method *sep-rho*. Here, we implement the *sep-rho* method as the $\rho$-setting policy.

**Fixing variables** A common heuristic in the PHA literature is to fix variables when there has been agreement across scenarios for a number of iterations. This can significantly reduce the time to convergence, especially in the presence of binary or integer variables. The caveat associated with this heuristic is that it may lead to suboptimal solutions. In our case, this is attenuated by the availability of optimality gaps during algorithm execution. This heuristic is already included in mpi-sppy, so its implementation, as well as experimentation with different values for the number of iterations that triggers the fixing, is greatly simplified. In our problem, the largest impact was observed when fixing investment variables, in particular transmission lines, after agreement around the lower bound, i.e. eliminating candidates that are not built in any of the scenarios considered.

### C. mpi-sppy spokes

Another important feature of mpi-sppy, in addition to its parallel implementation of PHA and related features and heuristics, is the hub-and-spokes architecture. In the mpi-sppy architecture, processes[2] are sorted into different functional groups, called *cylinders*. The main cylinder is the *hub*, which solves the conventional PHA, with subproblem solutions being parallelized across the processes that belong to it. A known shortcoming of the PHA is that before convergence, there is no guarantee regarding the quality of the solution. Furthermore, for nonconvex problems (like MILPs) or if heuristics like variable fixing are used, even convergence brings no such guarantees. In mpi-sppy, optimality bounds are obtained by the other groups of processes, called *spokes*. Each spoke runs an algorithm that can be decomposed by scenario subproblems and parallelized across the spoke's processes. They can get information from the hub to form these subproblems, and return an outer or inner bound of the stochastic optimization problem back to the hub. Outer and inner bounds[3] are used by the hub to obtain optimality gaps and decide on an early termination of the overall algorithm.

An example of an inner bound spoke is successively evaluating the first-stage solution of each subproblem on all other subproblems to obtain a fully implementable (feasible) solution. An example of an outer bound spoke is running a variant of PHA without quadratic terms, which under certain conditions provides valid bounds of the stochastic problem [18]. Several spokes of each type may be run concurrently, so that the best bound obtained so far can be used by the hub.

The mpi-sppy package includes various types of inner and outer bound spokes. Several combinations were tested, with satisfactory results being obtained when using the *Lagrangian* outer bound spoke and *looper* inner bound spoke. Our tests have not been exhaustive, so other spoke combinations may behave better as more scenarios or other power systems are considered in the continuation of this work.

## IV. Computational study

At the current stage of our research, we are interested in verifying that we can solve a large instance of our model by the method described in the previous section in a reasonable time. To do that, we implement the stochastic optimization model on a large, realistic power system and vary the number of scenarios included.

### A. Test case and scenarios

**CATS system** We test our model on an extension of the *CAlifornia Test System* (CATS) developed in [2], adapted for capacity expansion planning. The CATS is a geographically realistic representation of the California generation and transmission system, which includes over 8,000 buses and 10,000 transmission branches, with locations based on publicly available data and simulated, but realistic values for electrical parameters that are not publicly available.

**RPS target** Current California legislation states an RPS target of 60% by 2035, with a goal of 100% renewable energy by 2045 [19]. With current technology, such a target, if not interpreted as a *net* target, can only be obtained at prohibitively expensive costs. For the tests conducted here, an intermediate but still ambitious RPS target of 70% by 2045 was chosen.

**Climate-informed scenarios** Load and generation profiles used are generated based on projected weather and climate data from a regional refinement for California of the E3SM global circulation model for the year 2045, using a 3km-resolution grid [12]. A probability space can be constructed from the projected data, where each element corresponds to one day of hourly data for the target year. The space is reduced to a sample space via clustering using the *k-means* algorithm on a normalized feature vector of load and generation profiles using the default implementation of *k-means* available in *scikit* [20]. This process is similar to that of [8], but using projected

[2]Note that each process may, and in our case, does, have access to several CPU cores.

[3]For minimization problems, outer bounds are lower bounds and inner bounds are upper bounds.

weather data rather than historical data. The size of the sample space is varied in different tests cases. Each sample in the sample space corresponds to a scenario of the stochastic optimization problem.

**Algorithm and computational parameters** All tests were done on the *quartz* HPC cluster at Lawrence Livermore National Laboratory. Each node of the cluster has 36 2.1GHz Intel Xeon cores and 128GB RAM. Within each cylinder, 3 CPU cores are dedicated to each optimization subproblem, with each subproblem corresponding to a scenario of the stochastic optimization problem. Three cylinders were used for each test: one hub, one lower-bounder and one upper-bounder, so that each test instance uses 9 CPU cores per scenario. Our model was implemented using mpi-sppy 0.11.1.2 and *Pyomo* 6.2. Gurobi 10.0.2 was used to solve the MILPs. To simultaneously avoid spending too much time on one individual subproblem and avoid sacrificing MILP quality at later PHA iterations, a time-dependent MILP gap scheme was implemented, where the MIP gap was increased from 1% to 2% after 100s of subproblem solver execution, 5% after 200s, and 10% after 300s. For all instances, the PHA iteration limit was set at 60.

The different techniques described here that have been used to solve this problem introduce numerous parameters, and hence degrees of freedom, into our algorithm. For the purpose of this study, parameters have only been varied from their default values to the extent necessary to obtain a solution within the time limit imposed. Although there is room for further exploration of the interaction between solution time and solution quality across test instances, fixed parameter settings across test instances were found to be sufficient to achieve the desired MIP gap within the time limit imposed.

### *B. Results and Discussion*

Table I summarizes computational results for all instances tested. As expected for problems of this size, all instances hit the iteration limit before PHA convergence. In all cases, a solution with an optimality gap below 5% is obtained in under 4h. This gap is in the vicinity of 2% for the cases with over 100 scenarios, which may be of more interest to decision makers. These results suggest that when maintaining a constant ratio of cores per scenario subproblem, the growth in total computation time as more scenarios are included remains tractable.

We expect that accounting for the uncertainty in climate projections will require including many more scenarios in our pool. Nevertheless, we find these results encouraging for two reasons. First, they suggest that reasonably low computation times can be obtained by keeping the ratio of cores per scenario subproblem constant. It is worth pointing out that our largest scenario instance only used about 3% of the computational capacity of the HPC cluster, so a more than tenfold increase in the number of scenarios seems well within reach. Second, scenario reduction techniques are considered outside of the scope of this paper, but it is clear that as many more scenarios are included, some scenario reduction methodology will be necessary. In [8], with computational limitations in mind, the initial space of samples is reduced to the order of 50 scenarios. We expect that being able to handle a much larger number of scenarios than that, together with improved scenario reduction techniques, will allow working with a larger climate uncertainty set.

As mentioned before, one important feature of mpi-sppy is providing optimality bounds during the excution of the PHA. The evolution of the bounds for each case is shown in Fig. 2. All tests were run for 60 iterations, but the progression of optimality bounds shows that, if desired, termination could have been triggered earlier by reaching a slightly looser optimality target. Moreover, these results suggest that, as is common for this type of problems, good-quality feasible solutions can be obtained relatively quickly, within 10 or 20 iterations. After that, most of the gap reduction is provided by tightening lower bounds. In our tests, computational resources were allocated symmetrically to all cylinders. Some improvement could likely be obtained by dedicating more computational power to the more demanding, but more impactful, outer-bounding cylinder, especially after the initial iterations are finished.

TABLE I
OPTIMALITY GAPS ACHIEVED, TOTAL COMPUTATION TIME AND NUMBER OF NODES DEDICATED, FOR VARIOUS NUMBERS OF SCENARIOS.

| Number of scenarios | CPU Cores | Wall-clock time (min) | PHA Iterations | Optimality gap achieved (%) |
|---|---|---|---|---|
| 10 | 90 | 85 | 60 | 4.8 |
| 20 | 180 | 102 | 60 | 4.1 |
| 50 | 450 | 132 | 60 | 3.8 |
| 100 | 900 | 170 | 60 | 2.1 |
| 200 | 1,800 | 200 | 60 | 2.8 |
| 360 | 3,240 | 206 | 60 | 2.4 |

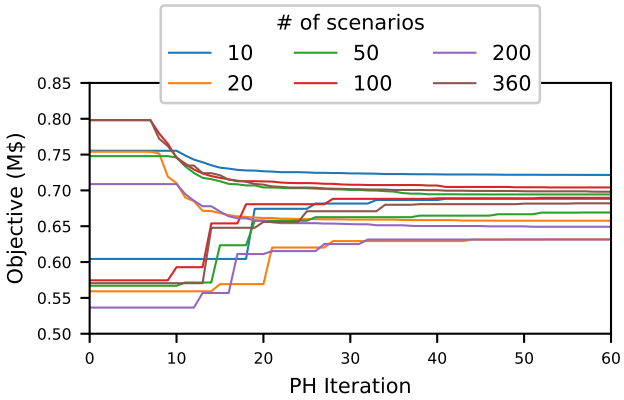


Fig. 2. Optimality bounds provided by mpi-sppy for different number of scenarios during the execution of the algorithm.

In Fig. 3, we compare the optimal capacity expansion plan across instances tested. It is interesting to note that as more scenarios are considered, both the optimal cost and the optimal investment plan continue to vary. This behavior could perhaps be mitigated by more sophisticated scenario selection methods, but nonetheless highlights the importance of using a sufficiently large scenario set in the stochastic program.

It is also noteworthy that no storage facilities were built in any of our test cases, although it is imaginable that this could change if the RPS target were increased. While our focus in this paper is on the computational aspects of our method and we leave a thorough analysis of the obtained results for future work, it seems clear that at least part of the value of storage may fail to be captured in the current model. This could be improved by including planned retirements of the existing generation fleet, as well as model modifications that highlight the value of flexibility brought by storage units: for example unit commitment and ramping, representation of security constraints, and improved power flow representation. These are all forthcoming extensions to the current model.

**RPS penalty and achieved RPS** A sensitivity analysis on the value of the soft constraint penalty $\lambda^{RPS}$ in (3) was performed using the case with 20 scenarios (Fig. 4). A value around \$80/MWh resulted in produced renewable energy near the target of 70%, so this was the value used for all the other test cases. This result appears to be relatively stable as the number of scenarios vary, as can be seen in Fig. 3, although the investments required to maintain it do change.

**Spurious losses** It is worth discussing why spurious transmission losses are a problem but spurious storage losses from simultaneous charging and discharging do not seem to be. This difference can be understood by comparing the *total* potential spurious losses and the *available* potential spurious losses of each type. By total potential spurious losses we refer to the total losses incurred if all resources were used at full capacity[4]: in our test system, this is in the order of 200MW for storage facilities, and 16GW for transmission assets. So the potential for impact is orders of magnitude higher with transmission than storage. To illustrate the second factor, note that the negative LMPs that drive spurious losses also encourage using storage facilities to store energy for future use (in a genuine, non-spurious way). Thus, using storage facilities to create spurious losses only makes sense if the storage facility is already full, which reduces the potential for spurious storage losses. This does not happen for transmission elements. The combination of these two factors explains why considering transmission losses has such a high impact on the relevance of spurious losses on our model.

As stated in section II-D, the change in the RPS constraint eliminates the induced negative LMPs, other than those potentially caused by congestion. As expected, spurious losses associated with these were found to be negligible in all tests.

## V. Conclusions

We have extended the stochastic capacity expansion model of [7], with the goal of representing a power system of sufficiently high-resolution to adequately capture climate variability and geographically realistic generation, storage, and transmission expansion decisions. To handle the high-resolution geographical and temporal data required for this improved accuracy, we implemented a variant of the model that could be decomposed by scenario and solved in an HPC environment using an implementation of PHA specifically tuned for our problem. mpi-sppy was used to facilitate the parallelization of the algorithm, as well as for obtaining optimality bounds during algorithm execution. Tests were conducted on a full-resolution, synthetic but realistic representation of the California generation and transmission system, with scenarios based on a high-resolution climate projection for 2045. Instances with up to 360 representative days were satisfactorily solved to under 5% optimality gaps within a few hours.

We envision the continuation of this research along two main directions: validation and improvement of the model and improvement of the computational solution. On the former, a sensitivity analysis of the impact of key model parameters and modeling assumptions (e.g. the chosen representation of power flow and the presence of resource flexibility constraints) on the obtained investment plan needs to be better understood. On the latter, our tests suggest that a significant performance improvement can be achieved by adopting an asymmetric management of computational resources across cylinders. In addition, addressing the difference in computational time among scenarios by way of an asynchronous decomposition of the optimization problem through scenarios, could have a significant impact on both solution time and quality. This is a functionality that has been recently added to mpi-sppy [16] and would constitute a natural extension of this work in the short term.

[4]This is not a feasible operation status, but serves as an upper bound for the total losses in the system, and hence for spurious ones.

## References


[1] J. D. Jenkins, E. N. Mayfield, E. D. Larson, S. W. Pacala, and C. Greig, "Mission net-zero America: The nation-building path to a prosperous, net-zero emissions economy," *Joule*, vol. 5, no. 11, pp. 2755–2761, Nov. 2021, publisher: Elsevier.

[2] S. Taylor, A. Rangarajan, N. Rhodes, J. Snodgrass, B. Lesieutre, and L. A. Roald, "California test system (cats): A geographically accurate test system based on the california grid," *IEEE Transactions on Energy Markets, Policy and Regulation*, pp. 1–12, 2023.

[3] H. Yuan, R. S. Biswas, J. Tan, and Y. Zhang, "Developing a reduced 240-bus wecc dynamic model for frequency response study of high renewable integration," in *2020 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, 2020, pp. 1–5.

[4] E. Raycheva, J. B. Garrison, C. Schaffner, and G. Hug, "The Value of Flexibility in a Carbon Neutral Power System," Nov. 2022, arXiv:2211.13625 [math].

[5] G. Migliavacca, M. Rossi, D. Siface, M. Marzoli, H. Ergun, R. Rodríguez-Sánchez, M. Hanot, G. Leclerq, N. Amaro, A. Egorov, J. Gabrielski, B. Matthes, and A. Morch, "The Innovative FlexPlan Grid-Planning Methodology: How Storage and Flexible Resources Could Help in De-Bottlenecking the European System," *Energies*, vol. 14, no. 4, p. 1194, Jan. 2021, number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[6] F. D. Munoz, B. F. Hobbs, J. L. Ho, and S. Kasina, "An Engineering-Economic Approach to Transmission Planning Under Market and Regulatory Uncertainties: WECC Case Study," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 307–317, Jan. 2014.

[7] R. S. Go, F. D. Munoz, and J.-P. Watson, "Assessing the economic value of co-optimized grid-scale energy storage investments in supporting high renewable portfolio standards," *Applied Energy*, vol. 183, pp. 902–913, Dec. 2016.

[8] F. D. Munoz and J.-P. Watson, "A scalable solution framework for stochastic transmission and generation planning problems," *Computational Management Science*, vol. 12, no. 4, pp. 491–518, Oct. 2015.

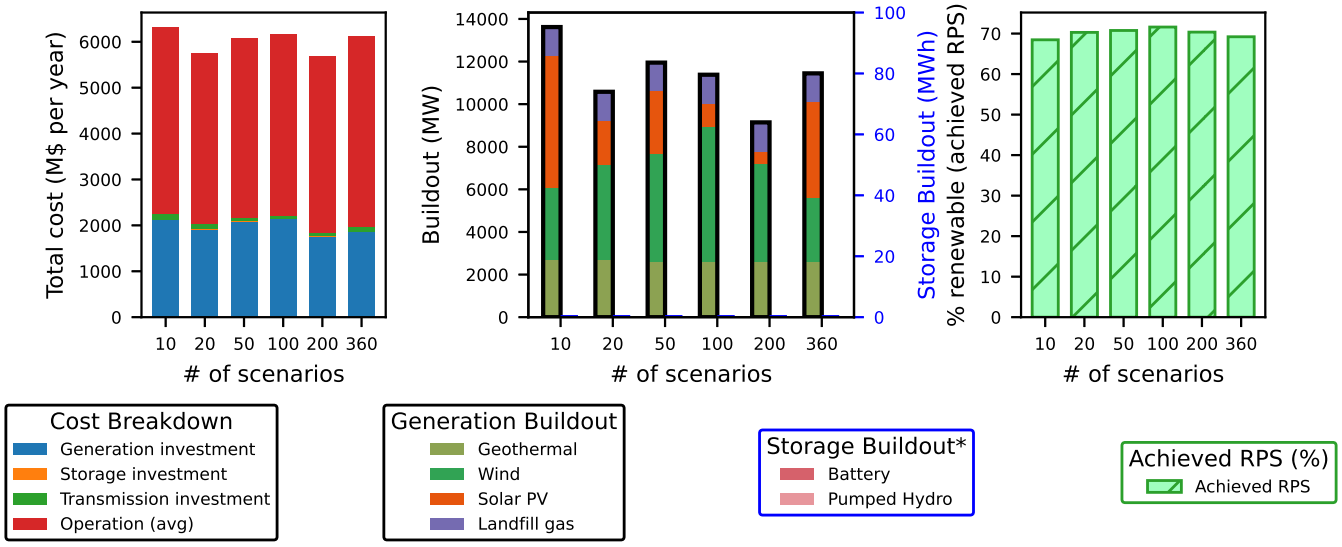


Fig. 3. Comparison of solution results for various numbers of scenarios. *Left*: Breakdown of optimal expected cost. *Center*: Generation and storage buildout. *Note that no storage was built in any of our test cases. *Right*: Expected proportion of energy generated with renewables.

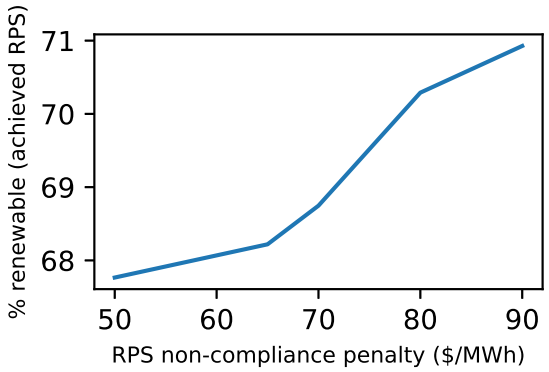


Fig. 4. Sensitivity of achieved proportion of total energy served with renewables to RPS penalty $\lambda^{RPS}$. Tests were performed using the instance with 20 scenarios.


[9] B. Knueven, D. Mildebrath, C. Muir, J. D. Siirola, J.-P. Watson, and D. L. Woodruff, "A parallel hub-and-spoke system for large-scale scenario-based optimization under uncertainty," *Math. Prog. Comp.*, Aug. 2023.

[10] M. L. Bynum, G. A. Hackebeil, W. E. Hart, C. D. Laird, B. L. Nicholson, J. D. Siirola, J.-P. Watson, and D. L. Woodruff, *Pyomo–optimization modeling in python*, 3rd ed. Springer Science & Business Media, 2021, vol. 67.

[11] W. E. Hart, J.-P. Watson, and D. L. Woodruff, "Pyomo: modeling and solving mathematical programs in python," *Mathematical Programming Computation*, vol. 3, no. 3, pp. 219–260, 2011.

[12] J. Zhang, P. Bogenschutz, Q. Tang, P. Cameron-smith, and C. Zhang, "Leveraging Regional Mesh Refinement to Simulate Future Climate Projections for California Using the Simplified Convection Permitting E3SM Atmosphere Model Version 0," *EGUsphere*, pp. 1–49, Oct. 2023, publisher: Copernicus GmbH. [Online]. Available: https://egusphere.copernicus.org/preprints/2023/egusphere-2023-1989/

[13] F. Neumann, V. Hagenmeyer, and T. Brown, "Assessments of linear power flow and transmission loss approximations in coordinated capacity expansion problems," *Applied Energy*, vol. 314, p. 118859, May 2022.

[14] D. Z. Fitiwi, L. Olmos, M. Rivier, F. de Cuadra, and I. J. Pérez-Arriaga, "Finding a representative network losses model for large-scale transmission expansion planning with renewable energy sources," *Energy*, vol. 101, pp. 343–358, Apr. 2016.

[15] J.-P. Watson and D. L. Woodruff, "Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems," *Computational Management Science*, vol. 8, pp. 355–370, 2011.

[16] J. Eckstein, J.-P. Watson, and D. L. Woodruff, "Projective hedging algorithms for multistage stochastic programming, supporting distributed and asynchronous implementation," *Operations Research*, 2023.

[17] J.-P. Watson and D. L. Woodruff, "Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems," *Comput Manag Sci*, vol. 8, no. 4, pp. 355–370, Nov. 2011.

[18] D. Gade, G. Hackebeil, S. M. Ryan, J.-P. Watson, R. J.-B. Wets, and D. L. Woodruff, "Obtaining lower bounds from the progressive hedging algorithm for stochastic mixed-integer programs," *Math. Program.*, vol. 157, no. 1, pp. 47–67, May 2016.

[19] California Public Utilities Commission, "Renewables Portfolio Standard (RPS) Program." [Online]. Available: https://www.cpuc.ca.gov/rps/

[20] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.