Energy Strategy Reviews 38 (2021) 100755

Contents lists available at ScienceDirect

Energy Strategy Reviews

journal homepage: www.elsevier.com/locate/esr

# Solving a large energy system optimization model using an open-source solver

Madeline Macmillan [a,b,*], Kelly Eurek [b], Wesley Cole [b], Morgan D. Bazilian [c]

[a] *Advanced Energy Systems Graduate Program, The Colorado School of Mines, Colorado, United States*
[b] *National Renewable Energy Laboratory, Colorado, United States*
[c] *Payne Institute, The Colorado School of Mines, Colorado, United States*

ARTICLE INFO

*Keywords:*
Open-source solvers
Energy models
Solve time
ReEDS

ABSTRACT

Open-source energy models are becoming more widely used for electric power systems planning. The solutions for these models are often computed using commercial optimization solvers, which require licensing fees that can be a potential barrier for certain organizations and researchers. This study explores the ability of the open-source COIN-OR linear programming (CLP) solver to compute solutions for the Regional Energy Deployment System (ReEDS) model—a large-scale, open-access electricity system planning model for the United States developed by the National Renewable Energy Laboratory (NREL). We find that open-source solvers, such as CLP, require some reduction of model size and detail. Although the solutions for reduced-form models may differ from full-featured models, we demonstrate that reduced-form solutions for ReEDS can still provide useful insight about drivers of power sector evolution. This research can help the modeling community better understand how open-source solvers can be applied to large-scale planning tools, and the potential steps that may be required to implement them.

## 1. Introduction

Energy system models are essential for informing researchers, utility owners, and policy makers. These tools vary by, among other aspects, sectoral coverage, spatial and temporal scales, and application. Until recently, most energy-system models available—even those of which the underlying code was made freely available—required licensing of commercial solvers. In the case of the Network-Enabled Optimization Server (NEOS) server, an open-source server with several solvers available, anyone can access the tool, but only for non-commercial purpose. More recently, a handful of research organizations have developed similar, more accessible tools that provide not just the model code freely, but were also developed to use open-source or open-access solvers [1–3].

The Open-Source Initiative, which aims to promote and protect open-source software, defines open-source software as that which can be developed and distributed by many people in a way that "grants all the rights to use, study, change, and share the software in modified and unmodified form" [4]. The essential component of open-source software is to "enable community development". Within open-source software initiatives, there are efforts focused specifically on energy system modeling. Open energy modeling applies the same principles as open software: "source code that can be studied, changed, and improved as well as freely available energy system data" [5]. Regardless of the application of "open", the idea to provide accessible products and platforms for collaborative development for all remains consistent [6]. In this study, we consider open-source solvers.

There are many advantages to open-source resources and software. Providing access to a tool can increase the sustainability of the project, promote collaboration, and further development of the tool. There are also possible disadvantages, including the possibility of low-quality end-user documentation, the constant evolution of software with no guarantee of cohesiveness across versions, potential security violations, and loss of competitive advantage [6]. Another potential drawback of

*Abbreviations:* CLP, COIN-OR linear programming; ReEDS, Regional Energy Deployment System; NREL, National Renewable Energy Laboratory; NEOS, Network-Enabled Optimization Server; GAMS, General Algebraic Modeling System; LP, linear program; ERCOT, Electric Reliability Council of Texas; MPS, mathematical programming system; BDMLP, Brook, Drud, and Meeraus Linear Program solver; GLPK, GNU Linear Programming Kit; IPOPT, Interior Point Optimizer; MINOS, Modular In-core Nonlinear Optimization System; PIPS, Parallel Interior Points Solver; SOPLEX, Sequential Object-oriented simPlex.

* Corresponding author.
*E-mail address:* madeline.macmillan@nrel.gov (M. Macmillan).

https://doi.org/10.1016/j.esr.2021.100755
Received 31 July 2021; Received in revised form 21 October 2021; Accepted 8 November 2021
Available online 25 November 2021

open-source or open-access software, particularly open-source solvers, are their solving capabilities relative to similar, commercial alternatives.

Optimization solver software is essential to computing solutions for energy system modeling tools. A barrier to entry of some commercial software is the licensing cost, although some commercial solvers are available for free or at low cost to academic communities [7]. Organizations or researchers with limited budgets, however, may not be able to afford commercial solvers. Therefore, it is of interest to identify open-source solvers that can produce robust results. Past efforts have evaluated open-source solvers for linear programming problems. A study from 2008 applies three different solvers for electricity spot market problems [8]. Another study from 2013 applies the same three solvers for a cell suppression problem [9]. The fastest solver is different for each study due to the inherent difference in the problem structures.

In this paper, we compare the capability of open-source and commercial linear programming solvers to compute solutions for a large-scale energy model. We evaluate solver performance when applied to problem instances of the Regional Energy Deployment System (ReEDS) model developed by the National Renewable Energy Laboratory (NREL) [10]. First, we introduce the ReEDS model (Section 2). We then compare the merits of several open-source linear programming solvers to select a candidate to apply to ReEDS and explore methods for improving the solvability and speed of ReEDS (Section 3). Next, we summarize the performance of an open-source solver applied to several reduced-form versions of ReEDS and examine the model outputs (Section 4). Finally, we propose next steps for future research (Section 5).

## 2. Model background

As previously mentioned, "open-source" is a form of software available for development and distribution by any interested party. Although ReEDS is available for distribution through an NREL form, users are not permitted to redistribute to third parties [11]. To access the ReEDS repository, follow the link in SI-1. The usage of ReEDS also requires paid licenses such as General Algebraic Modeling System (GAMS). This means that although ReEDS does not fit the definition of "open-source" it does have a level of openness.

The ReEDS model is a long-term planning model for the electric power sector. Given assumptions about future conditions (e.g., technology costs and performance, fuel prices, policy), the model determines the least-cost mix of generation, transmission, and storage resources necessary to meet physical constraints and policy requirements. The model is populated by several parameters including annual capital expenditures, levelized costs of energy, and capacity factors, many of which are housed within the Annual Technology Baseline [12]. The model was first developed for the conterminous United States but has since been extended to Canada, Mexico, and India [13–15]. The core of ReEDS is a linear program (LP) that minimizes the net present value of electric power sector costs subject to a suite of constraints governing the investment and operation of supply-side resources on a substate resolution. The constraints include balancing supply and demand for electricity, meeting reliability requirements for planning and operating reserves, abiding by physical operational constraints, transmission flows, and compliance with state and federal policies.

ReEDS is typically solved myopically with limited foresight about future conditions to inform investment decisions. As the model steps forward in time, new information about the future is revealed, and new decisions are made. In this sequential solve procedure, investments made in prior years affect decisions in future years.

ReEDS has a modular structure. Information is passed between modules to optimize or calculate various outcomes of the long-term planning process. A sequential ReEDS model solve begins with a supply module being provided inputs from previous model years. The solution and outputs are passed to a variable renewable energy and storage module to calculate the capacity values and curtailment rates of variable renewable generation technologies and storage technologies using hourly chronological data. These values are then given to the supply module to solve the next model year. This process continues until the end of the time horizon is reached (which is typically 2050) [10].

Another characteristic of ReEDS is how it treats historical years differently than present and future years. For example, new investments are limited to the exogenous capacity prescriptions in the years 2010–2018. Endogenous investments and endogenous retirements are enabled in 2020 and 2024, respectively. These years are updated as current years become historical years. Because historical plants are tracked separately from new plants, this results in an increase in the number of generation and storage resources within the model results and increases the size of the model A-matrix (number of rows, columns, and nonzeros). The ReEDS model has been used by various researchers for a variety of research questions [16–20]. Therefore, ReEDS is a valuable starting point to understand the ability of open-source solvers to solve the linear program.

## 3. Methods

### *3.1. Evaluation of linear programing solvers for ReEDS*

With a basic understanding of the ReEDS model, we began looking at different linear programming solvers to identify the best one to test. From literature, we compiled a list of candidate open-source solvers for ReEDS, identified available linear program solution algorithms with each solver, classified the solvers' compatibility with ReEDS, and examined their solve time performance for standardized model instances from benchmark studies. Based on these criteria, we selected one candidate open-source solver to test on ReEDS.

The ReEDS linear program is written in the GAMS mathematical programming software and is typically solved using the CPLEX commercial solver, applying the interior point method plus crossover to obtain a basic feasible solution. Therefore, candidate open-source solvers for ReEDS should be compatible with GAMS and should contain the interior point method.

In one 2020 study of energy system optimization model performance, Scholz et al. (2021) note that the interior point method usually outperforms both the primal and dual simplex methods for solving large-scale LPs of the electricity system [21]. Based on the experience of practitioners at NREL, the interior point method is faster than the simplex method for solving common ReEDS instances [22]. Klotz and Newman offer insights into the potential performance of these two methods based on the A-matrix characteristics of a difficult linear program problem instance [23].

For this study, we considered the following solvers as potential lower-cost alternatives to CPLEX for the ReEDS model:

- **Brook, Drud, and Meeraus Linear Program solver (BDMLP)** [24] – Linear programming solver managed by GAMS. The solver is not open source, but it is available with the purchase of a GAMS license. BDMLP has been dropped from the GAMS distribution as of GAMS version 34.
- **COIN-OR Linear Programming (CLP)** [25,26] – Free, open-source linear programming solver made available through the COIN-OR project. A link between GAMS to CLP is available with the purchase of a GAMS license.
- **GNU Linear Programming Kit (GLPK)** [27] – Free, open-source linear programming package. GLPK is no longer part of the GAMS distribution.
- **Interior Point Optimizer (IPOPT)** [28] – Free, open-source suite of interior point solvers for linear and nonlinear optimization problems available through the COIN-OR project. A link between GAMS and IPOPT is available with the purchase of a GAMS license.
- **LP Solve** [29] – Not commonly used for GAMS models, but examples of links between GAMS and LP Solve exist. There are examples of LP Solve being used for models developed in R programming language.

- **Modular In-core Nonlinear Optimization System (MINOS)** [30, 31] – Commercial software with discounted academic licenses and government licenses. A GAMS/MINOS-Link is not offered to GAMS customers, so an organization such as a government-based research institution, would have to pay the standard rate for a GAMS/MINOS license.
- **Parallel Interior Points Solver (PIPS)** [32] – Free, open-source suite of parallel optimization solvers developed by Argonne National Laboratory. A PIPS-IPM solver link for GAMS was developed for the BEAM-ME Project [33], but the link is not publicly available.
- **Sequential Object-oriented simPLEX (SOPLEX)** [34,35] – Commercial optimization package for linear programming problems. Free versions are available for noncommercial and academic institutions and links to GAMS are available with the purchase of a GAMS license.

Table 1 summarizes the criteria we used to evaluate the eight solvers listed above for use with ReEDS, but this table can also reflect solver compatibility with other programming platforms dependent on the energy model in question. We gave priority to solvers that: (1) are open-source, (2) include the interior point method—the recommended solution method for ReEDS is interior point with crossover, (3) are compatible with GAMS—the ReEDS formulation is written in GAMS, (4) performed well in past benchmark studies, and (5) are available at low or no cost—software fees are an important factor for organizations with limited budgets. The execution of the ReEDS model requires a fee-based license for GAMS, so the additional cost for a commercial solver like CPLEX increases the cost burden. If all other criteria are met, then we used the benchmark performance as a deciding factor for which solver may exhibit superior performance for ReEDS.

The benchmark performance metrics provide insights into how the solvers compare with each other. A 2013 report from Sandia National Laboratories compares the solvability and solve time of CPLEX and a variety of open-source solvers (CLP, GLPK, LP Solve, MINOS) on a suite of 180 standardized linear programing problem instances of varying sizes (rows, columns, and nonzeros) [38]. The largest problem tested on these solvers has more than 1.9 million rows and more than 0.64 million columns. For reference, this problem is larger than the largest ReEDS problem size after the CPLEX presolve that occurs in the final year of the modeling horizon, in this case the year 2050, which has 0.44 million rows and 0.52 million columns.

The authors find the CLP solver can solve the tested problems faster than any of the other open-source solvers tested, but CPLEX is superior across all test problem instances. Among the open-source solvers tested on the 180 problem instances, CLP was able to solve the most problems out of the open-source solvers tested. Within the Hans Mittleman "Benchmarks for Optimization Software," CLP generally yielded faster solve times compared to other open-source linear program solvers [39]. Finally, a 2006 study found that CLP outperformed GLPK and LP Solve for electricity spot market optimization problems [8]. While the electricity spot market problem does not consider investment decision-making, the problem includes many of the same types of operational variables and constraints as ReEDS.

Based on Table 1, we excluded solvers that:

- Are not open source (BDMLP, PIPS, and SOPLEX)
- Are not free or are not low-cost for nonacademic use (SOPLEX; MINOS)
- Do not use the interior point method (BDMLP; SOPLEX; MINOS; LP Solve)
- Are not directly accessible in GAMS (PIPS; GLPK; LP Solve)
- Are not tested in past benchmark studies (BDMLP; IPOPT; PIPS; SOPLEX).

The CLP solver is the only open-source solver to satisfy all our criteria for use with the ReEDS model and was therefore selected for further characterization. Although IPOPT is not represented in the benchmark studies, it satisfies every other requirement. In an initial test, we found IPOPT produces solutions that are not compatible with the ReEDS sequential-solve algorithmic structure [11]. For example, after solving one model year, IPOPT would include nonzero values for variables not considered in the model. These nonzero values would create issues when

**Table 1**
Candidate solver options for ReEDS.

| Solver | Open Source? | Solver Cost [29,30,36] | LP Solution Algorithms | Compatibility with GAMS [37] | Performance benchmarks in [38] | Performance benchmarks in [39] |
|---|---|---|---|---|---|---|
| BDMLP | No | Only available through commercial software | • Simplex | Managed by GAMS and available with the GAMS Base Module License | Not included | Not included |
| CLP | Yes | Free | • Simplex<br>• Interior Point | GAMS solver link available with the GAMS Base Module License | • Successfully solved 180/180 problems<br>• Aggregate solve time 213 times slower than CPLEX | Successfully solved 40/40 problems |
| GLPK | Yes | Free | • Simplex<br>• Interior Point | No longer part of the GAMS distribution | • Successfully solved 138/180 problems<br>• Aggregate solve time 459 times slower than CPLEX | Successfully solved 36/40 problems |
| IPOPT | Yes | Free | • Interior Point<br>• (no crossover) | GAMS solver link available with the GAMS Base Module License | Not included | Not included |
| LP Solve | Yes | Free | • Simplex | Not commonly used for GAMS models and no formal GAMS solver link is available | • Successfully solved 150/180 problems<br>• Aggregate solve time 510 times slower than CPLEX | Not included |
| MINOS | No | • $936 – academic[a] and government<br>• $25,000 – commercial[b] | • Simplex | GAMS solver link does not exist, accessible with a GAMS\MINOS license | • Successfully solved 151/180 problems<br>• Aggregate solve time 613 times slower than CPLEX | Not included |
| PIPS | Yes | Free | • Simplex<br>• Interior Point | Open-source solver, GAMS solver link is not publicly available. | Not included | Not included |
| SOPLEX | No | • Free – academic<br>• High – commercial | • Simplex | GAMS solver link: available with the GAMS Base Module License (Academic Only), also accessible with a GAMS/SCIP license | Not included | Successfully solved 36/40 problems |

[a] University-wide license.
[b] Company-wide license.

solving later model years. This is explained in greater detail in SI-2.

### 3.2. Improving the solvability of the ReEDS reference case on CLP

We tested CLP on the ReEDS "reference case" for two spatial extents: (1) the Electric Reliability Council of Texas (ERCOT); and (2) the conterminous United States (CONUS). CLP was able to solve the ERCOT model instance but failed to solve the CONUS model instance for the full model planning horizon within the 10,000 s (~2 h 47 min), our default solver time limit for this study.

Excessive run time can be indicative of a large problem size that is cumbersome for the solver and/or a problem instance with numerical issues [23]. Numerical issues, such as inaccuracies from round off errors that occur with floating point calculations (ex. a poorly scaled A-matrix with coefficients that have a large difference in their orders of magnitude), will hinder the performance of any solver, but more so for free solvers that have less sophisticated algorithms to abate these numerical issues. Although CLP could theoretically solve a ReEDS model instance given sufficient time, it is not time efficient. It is also important to consider the impact the CPLEX presolve algorithms might have on model tractability when compared to those of CLP and how that may affect the solve time of the model.

By reducing the problem size or simplifying the model formulation of linear programs, solve time can be reduced [21]. Klotz and Newman as well as Scholz et al. offer guidelines for improving the solve times of linear programs [21,23]. Some of these guidelines are specific to CPLEX, while others are solver agnostic.

We identify several methods to improve the solvability of a large model on open-source solvers (Table 2). The following sections will provide more details on each technique employed for the ReEDS model in this study.

### 3.3. Improving the A-matrix

To identify possible areas for avoiding numerical issues, we inspect the A-matrix, b-vector, and c-vector through a mathematical programming system (MPS) file. All of these values are accessible through solver output. As a standardized format for storing linear programming problems, MPS files include the coefficients we aim to stabilize. MPS files are also organized by variable and constraint for easy classification of unstable constraints or areas for improvement. An example of the code we used to manipulate our MPS files can be found in SI-3.

#### 3.3.1. Round matrix coefficients

The first method attempted to improve the stability of the A-matrix was to round specific coefficients. Many of the coefficients, regardless of their scaling, had excessive and unrealistic precision (e.g., 40.0000001). Rounding should be performed with caution and should only be applied to input data. To reduce this source of instability, the ReEDS model code was edited to round the coefficients more effectively. An example of when this might be necessary and useful is when a 100-MW photovoltaic system with a capacity factor of 0.00005 produces 0.003 MW-hours of energy in an hour. In such a situation, it would be reasonable to round the capacity factor to zero and assume the PV system does not produce any energy during the period in question.

**Table 2**
Techniques to improve solvability of large models on open-source solvers.

| Category | Technique | ReEDS examples from this study |
|---|---|---|
| Reducing numerical issues in the A-matrix | Round matrix coefficients | Round emission rates |
| | Scale matrix coefficients | Scale emissions variables |
| Reducing the size of the A-matrix | Remove variables and constraints for: (a) low-impact features (b) advanced features | (a) Exclude variables and constraints for technologies that are unlikely to be deployed in a reference case (b) Do not allow endogenous retirements as a decision variable |
| | Reduce the model dimensions | Reduce the spatial extent of the model to ERCOT Reduce the number of investment periods |

#### 3.3.2. Scale matrix coefficients

The second method for improving stability was to adjust the scaling factors. Throughout the ReEDS formulation, there are instances of poorly scaled A-matrix coefficients. By improving the scaling of the A-matrix coefficients and moving them closer to unity, solve time can be reduced. The best practice is for the maximum difference between the smallest and largest coefficient orders of magnitude to be 12 (e.g., 1e-6 to 1e+6), but with smaller spreads resulting in better scaling [23]. By inspecting the matrix, we can round coefficients to reduce instances of very small coefficients, adjust scaling factors, and remove variables and/or constraints causing numerical issues while keeping in mind the impact on the solution.

There are a few options for scaling the coefficients in the A-matrix. The first is indicating scaling preferences within the solver options. For this study, both solvers with GAMS compatibility had a unique set of options [37]. Alternatively, manual scaling was possible with user-defined scaling factors. In this study, we adjusted the ReEDS scaling factor applied to emissions constraints because we identified them as a source of the poor scaling within the A-matrix. The ReEDS emissions constraints are responsible for maintaining emissions levels within the required emission limits. In the original problem formulation, the parameter is in units of megatons across the three pollutants considered in the model ($CO_2$, $SO_x$, and $NO_x$). Because power sector $CO_2$ emissions were orders of magnitude larger than $SO_x$ and $NO_x$ emissions, poor scaling ensued. To test whether scaling of the A-matrix coefficients might improve model tractability and solver solve time, we conducted a run where the emission scaling parameter was adjusted to be specific to each pollutant type [40].

### 3.4. Remove variables and constraints

The third method explored for improving solver performance was removing select sets of variables and constraints. Problematic variables and constraints, such as those with numerical issues, can be identified by filtering the MPS file, as necessary. We determined which variables and constraints to remove by (1) excluding features that may not have a strong impact on the solution, (2) excluding advanced features, and (3) reducing the model dimensionality. Within ReEDS, many of the relevant variables and constraints were those dictating emissions policies. These constraints, among a few others, were removed to understand if and how these changes would affect the solve time. A complete list of the various model features turned off throughout this study is shown in Table 3.

Model dimensionality is typically categorized by space, time, and the technologies represented. Across space, the number of regions is defined based on the spatial extent and the size of the regions. The spatial dimension can be reduced by limiting an analysis to a small number of regions and/or by clustering regions into larger groups [41]. For this study, we tested a scenario focused on the ERCOT interconnection, which represents about 5% of the total number of regions in the full U.S. ReEDS model.

In the time domain, models include both operation time periods (e.g., time slices) and investment time periods (e.g., years). Models need sufficient operational time periods to capture, for example, seasonal and diurnal patterns for load and renewable resource profiles. For investment time periods, some analyses focus on the year-to-year pathway from today to the future, whereas others focus on the system design for a single future year. For this study, we test a scenario with investment decisions made once in each decade versus the default of 2-year investment periods through 2030 and 5-year investment period thereafter.

With respect to technologies, models can track generation capacity

**Table 3**
Sample list of model features that can be turned off in ReEDS.

| Model Feature | Category | Description |
|---|---|---|
| RGGI (Regional Greenhouse Gas Initiative) | Emissions (EMIS) | Limit total $CO_2$ emission for states participating in RGGI |
| AB-32/SB-32 (Assembly Bill 32 and Senate Bill 32) | Emissions (EMIS) | California $CO_2$ cap and trade program |
| CSAPR (Cross State Air Pollution Rule) | Emissions (EMIS) | SOx/NOx emission caps for specific states |
| Endogenous retirements | Capital stock (CAP) | Endogenous decision for model plants to be retired prior to the end of their maximum lifetime |
| Capacity refurbishments | Capital stock (CAP) | Endogenous decision to refurbish a technology after the end of its lifetime |
| Carbon capture and storage | Technology (TECH) | Represent CCS technology options for coal and natural gas |
| State renewable portfolio standards (RPSs) | State RPS (RPS) | Enforce state-level RPSs, including constraints for renewable energy credit (REC) creation and REC trading |
| Operating reserves | Operating reserves (OR) | Balance the supply and demand for operating reserves and limit which technologies can provide reserves |

as individual generators or clusters of generators. Clusters are typically defined based on the location, age, and performance of the generators. For this study we did not adjust the default assumptions of thgrammare technology representation.

### 3.5. *Establishing a baseline for solve time*

Before testing the above methods, we established baseline solve times for both CPLEX (the default solver used by NREL [10]) and CLP. All runs were performed on an Intel(R) Xeon(R) Gold 5120 machine with CPU speeds of 2.20 GHz along with 14 Cores, 28 Logical Processors, and 768 GB of memory [42].

The default, U.S. reference case for ReEDS Version 2020, was used for the baseline [22]. Fig. 1 reports the baseline solve times for the two solvers. Table 4 summarizes the problem size before and after the CPLEX presolve. CPLEX employs a presolve that reduces the original size of the model. This is shown in the GAMS log as “reduced LP size”. The CLP output in the GAMS log was not available by default but was applied and did impact the size of the LP.

The recorded problem size was the same regardless of the solver used, so the values reported are from CPLEX output. Fig. 2 reports the model size as it increases throughout the solution time. Fig. 1, shows that the full ReEDS model can be solved by CPLEX, but after the year 2022, the solver CLP times out after the default timeout of 10,000 s spent on one model year. Fig. 2 visualizes the problem size after presolve and highlights how the ReEDS problem size increases throughout the model horizon. The baseline values in Fig. 2 are the same for both the commercial and open-source solver and were used throughout the study to characterize the various methods employed to improve tractability and solve time.

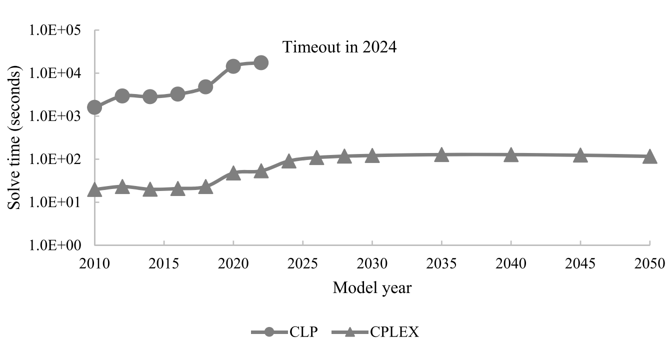


**Fig. 1.** CLP and CPLEX baseline solve times for the unaltered ReEDS model.

**Table 4**
Baseline problem size in 2050.

| ReEDS Model Instance | Rows | Columns | Nonzeros |
|---|---|---|---|
| U.S. (2050) – CPLEX | 3.0 million | 4.5 million | 23.3 million |
| U.S. (2050) – after CPLEX presolve | 0.44 million | 0.52 million | 2.4 million |

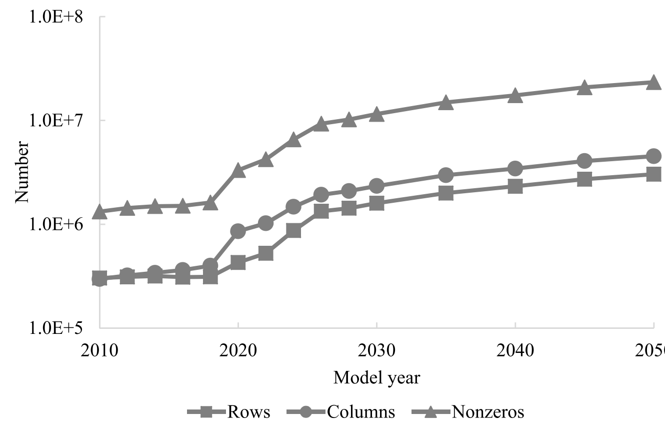


**Fig. 2.** Baseline problem size for full ReEDS model.

## 4. Results and discussion

Table 5 summarizes the suite of ReEDS scenarios that we attempted to solve using CLP. These scenarios are modified versions of the default formulation of the ReEDS model using strategies described in Section 2, including rounding parameters, scaling parameters, removing variables and constraints, and reducing the model dimensions.

### 4.1. *Exploring the solvability of ReEDS using CLP*

Fig. 3 summarizes the solve times associated with different scenarios using CLP. To compare the performance of CLP versus CPLEX, we include the solve times for BASE, ERCOT, and DECADES using CPLEX.

Fig. 3 shows that the removal of emissions constraints alone is not sufficient to enable CLP to solve additional model years within the cutoff time. The two scenarios with parameter modifications—Emit_rate and Emit_scale—were omitted from Fig. 3, because the modifications made to the model for these scenarios did not change the outcome relative to BASE. CLP was able to solve all model years (to 2050) for scenarios that removed combinations of variables and constraints, including emissions policies, capital stocks, RPS policies, reliability and carbon capture and storage (CAP TECH RPS OR, RPS OR, and ALL). The RPS OR scenario was solved by CLP for all model years with the fewest number of constraints turned off. This is important to note because fewer constraints turned off means fewer changes to the formulation, potentially resulting in the most comparable solutions to the BASE model. The implications of turning off these constraints on the ReEDS solution are explored in Section 3.2. The ALL scenario was solved by CLP for all model years with the fastest cumulative run time relative to other scenarios with variables and constraints removed. Although solve times for ReEDS scenarios using CLP were improved as more variables and constraints were removed, the removal of certain variables and constraints were more effective at accomplishing this goal.

To understand the source of solve time improvement, we reviewed changes in problem size. In the ALL scenario, the number of rows, columns, and nonzeros were all reduced from the full model, as seen in

**Table 5**
Summary of the ReEDS scenarios attempted using the CLP solver.

| Scenario Name | Technique | Description | Impact on 2010 solve time relative to CLP – BASE[a] | Percent change in objective value[b] |
|---|---|---|---|---|
| BASE | N/A | Default U.S. model | Timeout in 2024 | N/A |
| Emit_rate | Rounding parameters | Set the rounding of the emissions rate parameter to four decimal points | No solve time improvement through 2022 Timeout in 2024 | N/A |
| Emit_scale | Scaling parameters | Scale the emissions variables and constraints using a scaling parameter that is specific to the pollutant type | No solve time improvement through 2022 Timeout in 2024 | N/A |
| EMIS | Removing variables and constraints | Turn off all constraints associated with emissions | Approximately the same problem size as BASE No solve time improvement through 2022 Timeout in 2024 | N/A |
| RPS OR | Removing variables and constraints | Turn off state RPS requirements and operating reserve requirements | 20x reduction in solve time through 2022 Solves through 2050 | - 0.44% |
| CAP TECH RPS OR | Removing variables and constraints | Turn off constraints for capital stock such as endogenous retirements, technologies, state RPSs, and operating reserves | 20x reduction in solve time through 2022 Solves through 2050 | +0.73% |
| ALL | Removing variables and constraints | Turn off all model features indicated in Table 3 | 27x reduction in solve time through 2022 Solves through 2050 | +0.56% |
| ERCOT | Reducing the model dimensions | Reduce the spatial extent to the ERCOT system | 30x reduction in solve time Solves through 2050 | 0% |
| DECADES | Reducing the model dimensions | Original U.S. model solved only for 2010, 2020, 2030, 2040, and 2050 | Limited solve time improvement Solves through 2020 | N/A |

[a] Because several runs did not run to completion, one model year was selected for comparison.
[b] The percent change in the total system objective cost from CPLEX – BASE in the year 2050.

Fig. 4. As a result of the iterative sequential solving process in ReEDS, discussed previously in Section 1.1, the size of the problem increases significantly as the model progresses to years further into the future. In general, it was observed that a reduction in problem size helped the problem solve faster.

### 4.1.1. Model output

When a model formulation is adjusted to improve solvability of the model, the modified formulation may yield different results from the full-featured model. Past efforts in electricity system capacity expansion modeling have compared the trade-offs of model resolution, solve time, and outcomes [43,44].

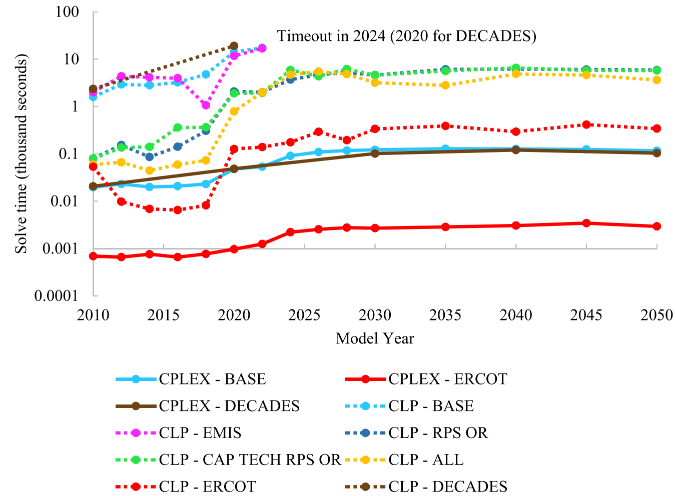


**Fig. 3.** Solve time summary.

Here we evaluate the impact of using reduced-form ReEDS problem instances on select model output metrics, including national $CO_2$ emissions, installed capacity, and system cost. We limit our inspection to scenarios that were solved by CLP for all model years (RPS OR; CAP TECH RPS OR; and ALL) and compare them to the BASE scenario solved by CPLEX (as the BASE scenario was not solved by CLP).

Fig. 5 shows the deviations in $CO_2$ emissions from the BASE scenario. The emissions deviations become more dramatic after 2022. This is in part due to how ReEDS treats historical years differently, as previously discussed. However, aside from that model characteristic, the RPS OR scenario deviates the least when compared to the other scenarios that also ran through 2050. A general observation can be made that when the renewable portfolio standards are omitted from the model scenarios, emissions increase significantly. This is confirmed by calculating the cumulative emissions over the entire modeling horizon and weighting each year to account for the step size (e.g., 2030 represents 2 years; 2045 represents 5 years). We find that the corresponding increases in cumulative $CO_2$ emissions from BASE for runs ALL, CAP TECH RPS OR, and RPS OR are approximately 5.9%, 5.7%, and 1.7%, respectively. This is an indicator that while CLP can solve a modified ReEDS scenario, it cannot solve high renewable energy and/or low carbon scenarios within 10,000 s per solve year on the machine tested.

To mitigate the effects of excluding the state RPS constraints, a practitioner could apply a renewable production incentive within the objective function to serve as a proxy for REC payments to renewable energy sources. Ultimately, practitioners must decide which sacrifices in model features are most appropriate given the analysis questions of interest. In a high-penetration renewable electricity future, state-specific RPS policies may no longer be binding. However, the regionality of RPS policies may have implications on renewable energy deployment in the near-term planning horizon, and thus the appropriate incentives should be captured in the model even if they are simplified.

Fig. 6 summarizes the deviations in national capacity from the BASE scenario for select technologies that experienced the most significant fluctuations, including coal, combined cycle natural gas, offshore wind, onshore wind, utility PV, and 4-h batteries. The RPS/OR scenario resulted in the least significant deviations except for the wind-offshore technology. For this technology, all scenarios responded with the same changes from BASE.

Fig. 7 summarizes the deviations in the objective function from the BASE scenario. The cumulative percent differences in system cost from BASE for ALL, CAP TECH RPS OR, and RPS OR are 1.31%, 1.22%, and −0.24%, respectively. The fuel cost increases are likely because of the increase in conventional generation in the system capacity due to the elimination of the RPS constraints. Across the different scenarios, the

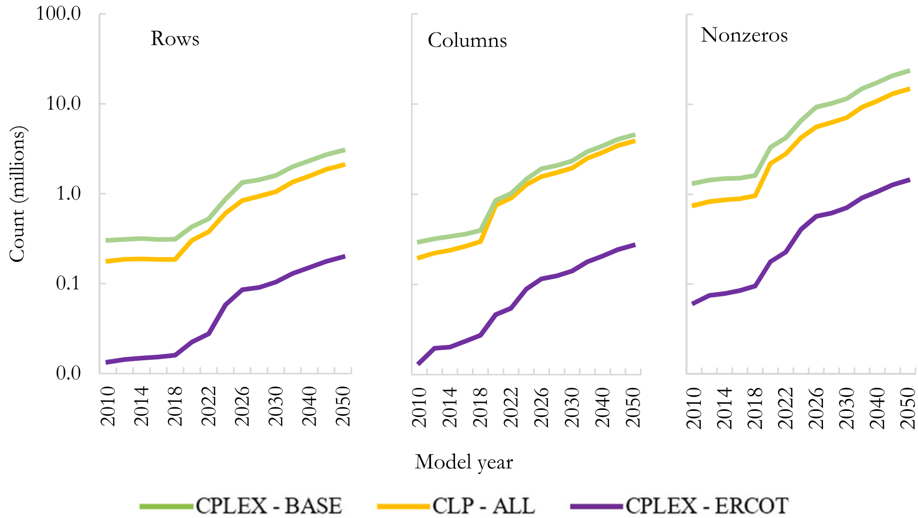


**Fig. 4.** Number of rows, columns, and nonzeros in the linear program ReEDS following formulation modifications.

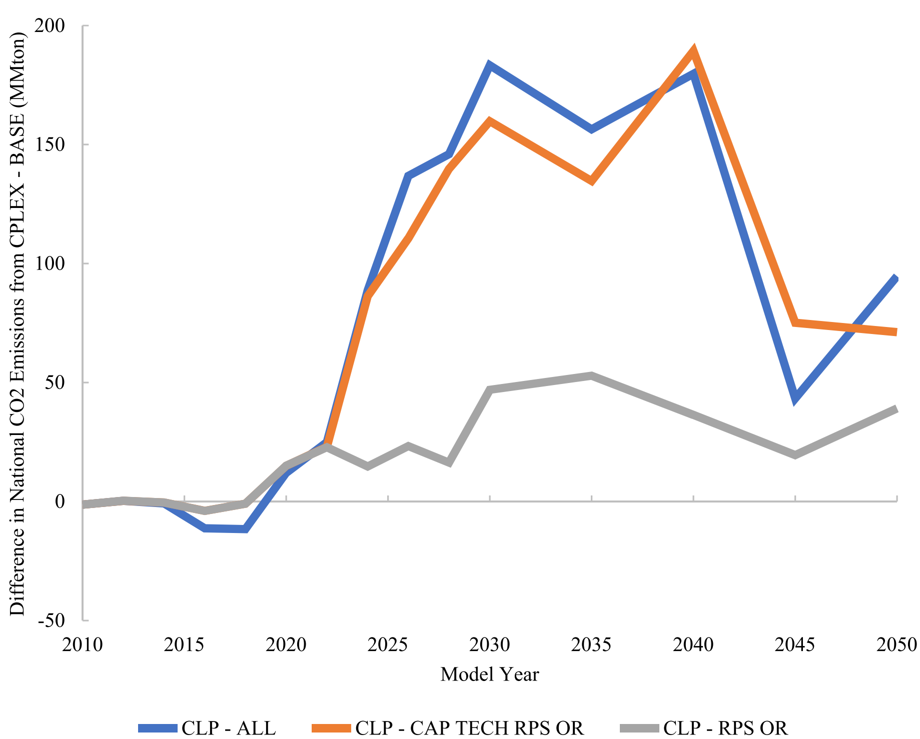


**Fig. 5.** Difference in national CO2 emissions from CPLEX – BASE, 2010–2050.

transmission cost experienced the most negative percent change from BASE. This can be explained by the fact that the model solution is relying on fewer renewable energy resources and more on centralized resources, therefore requiring less transmission infrastructure. Taken as an average across the different cost variables, the CAP TECH RPS OR scenario had the most significant percent change in magnitude from the BASE scenario.

The ReEDS modeling results are mostly prescriptive through the 2020 model year, as these are historical years, but we allow ReEDS to build combustion turbine gas technologies during these years as a slack variable for maintaining model feasibility in historical years.

The graphs in Fig. 5 through Fig. 7 highlight deviations from the BASE run across different metrics. The magnitude of the deviations became more substantial beyond the year 2022. The scenarios with fewer constraints eliminated (i.e., fewer model formulation modifications) resulted in less significant deviations from the BASE scenario. As modelers attempt to improve the solvability of their chosen model on open-source solvers, they should stay mindful of the modifications that result in the fewest constraint eliminations to produce results closets to the original model formulation.

## 5. Conclusions

In this study, we examine the potential for an open-source solver to compute solutions for a large-scale capacity expansion model of the U.S. electric power system. We identify the CLP solver as the most viable option to use for the ReEDS model based on several evaluation criteria, including: (1) it is free or low-cost access for all users, (2) it is open-

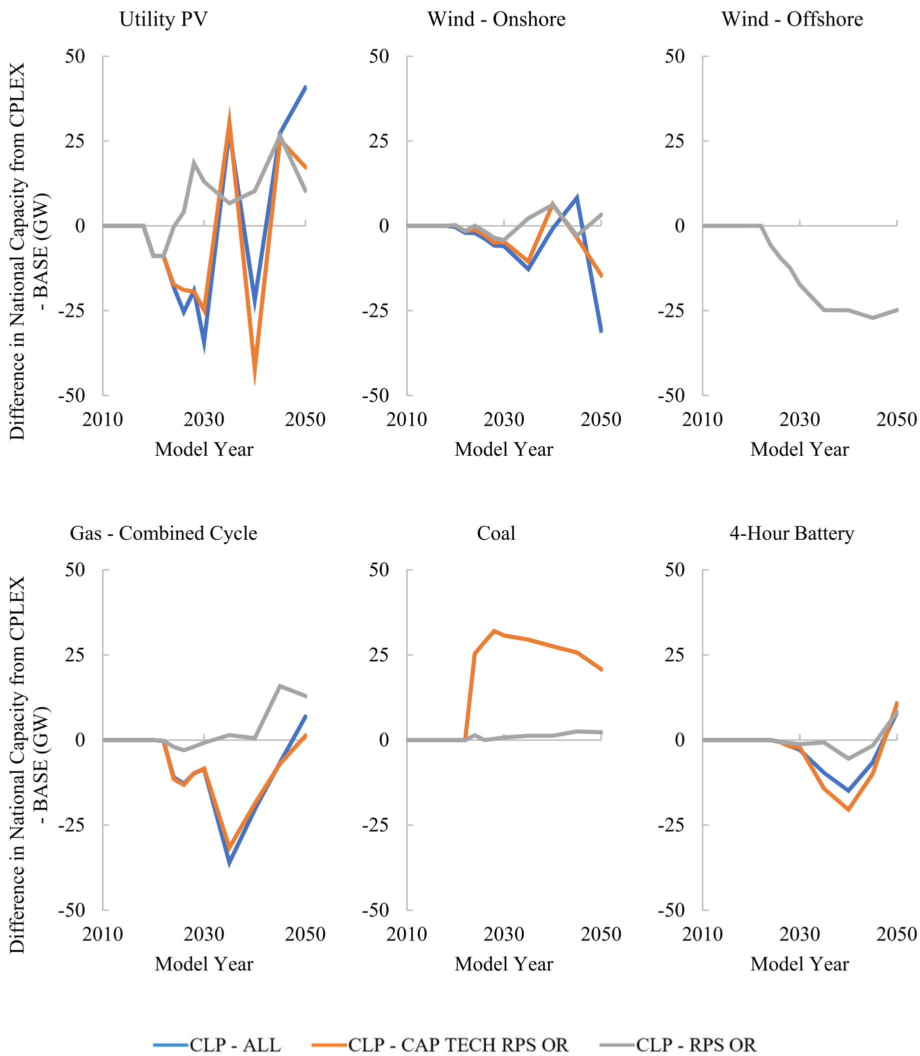


**Fig. 6.** Difference in national capacity from CPLEX – BASE by select technology, 2010–2050. (All three scenarios are depicted in each chart, however, in the instance of the coal graph, CLP – ALL is not visible. The CLP – ALL results align with those of the CLP – CAP TECH RPS OR run for coal. In the instance of the offshore wind graph, CLP – ALL and CLP – CAP TECH RPS OR results are not visible because they align with the results of the CLP – RPS OR result.)

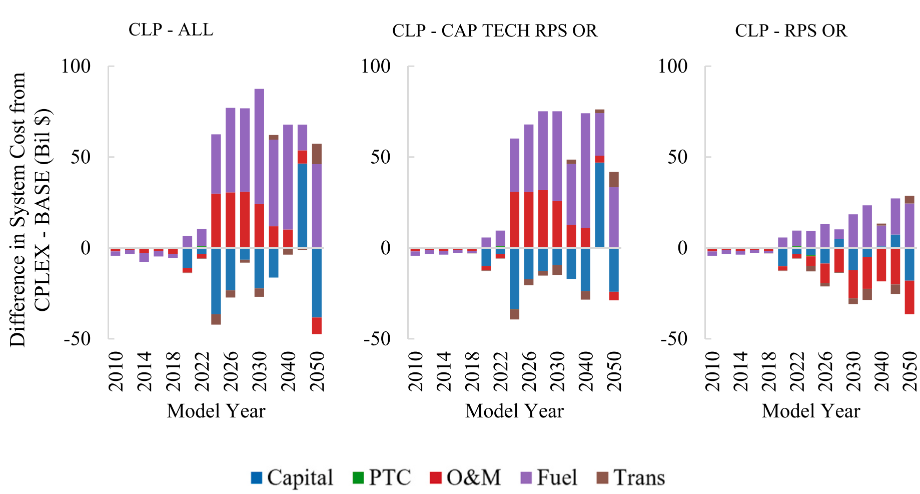


**Fig. 7.** Difference in system objective cost from CPLEX – BASE.

source or publicly available, (3) it includes the interior point method, (4) it is easily linked through GAMS, and (5) it is represented and performed well in past benchmark studies. Compared to other studies discussed in this paper, we conduct an analysis on solver performance for an energy model and determine methods for how the solver performance (i.e., solve time speed) can be improved through modifications to the energy model formulation. While we find CLP to be the best candidate open-source solver for ReEDS, CLP may not be the best option for all models. A limitation of our study is that only one model was tested. The ReEDS modeling community is not a complete representation of the energy system modeling community. Therefore, expanding this study to other large energy models would broaden the reach of this study's findings to more energy modelers. The techniques applied in this study to investigate the utility of open-source solvers on ReEDS can be used by other researchers on their own models of interest.

Although CLP was unable to solve the full-featured ReEDS model within the designated cutoff time of 10,000 s per solve year, it was able to solve reduced-form versions of ReEDS through the entire modeling horizon within the cutoff time. We reduce the problem size by excluding certain model features—regional emissions policies, endogenous retirements, carbon capture and storage technologies, state RPS policies, and operational reliability requirements—and reducing the dimensions of the problem. For the reduction of dimensions, we only explore a smaller spatial extent and fewer number of years modeled, but other dimensions could be reduced, including spatial resolution, temporal resolution, and technologies represented. For the runs with reduced spatial extent (i.e., ERCOT), CLP was able to solve ReEDS for all model years with all options turned on. Reducing the dimensionality allows practitioners to maintain the model features at the expense of lower-resolution model outputs. We compare the national-level emission, generation capacity, and objective function values of the reduced-form model scenarios with those of the full feature model and find noticeable differences in the results. The variation in the results will need to be considered by individual modelers in accordance with their project needs and goals. Some additional considerations for removing the constraints selected in this study are that a solution may not be compliant with state-specific renewable generation requirements without the state RPS constraint, and that a solution may not have sufficient capacity to meet unexpected changes in supply and demand without the OR constraint. Moving forward, since there will be instances of linear program open-source solvers struggling to solve very large problems, additional work to improve the solvers themselves can make them more useful to model practitioners.

Using open-source software to solve large energy system optimization models can increase the accessibility of energy model tools. This study offers techniques that will help enable the use of open-source solvers for energy system modeling for modelers regardless of their budget or available solver resources.

## Funding

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U. S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government.

## Credit author statement

Madeline Macmillan: Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision. Kelly Eurek: Conceptualization, Methodology, Software, Validation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. Wesley Cole: Project administration, Funding acquisition. Morgan D. Bazilian: Writing – review & editing, Supervision.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors would like to thank Daniel Steinberg, Alexandra Newman, and Patrick Brown for their assistance with the writing of this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.esr.2021.100755.

## References

[1] M. Groissböck, Are open source energy system optimization tools mature enough for serious use? Renew. Sustain. Energy Rev. 102 (Mar. 2019) 234–248, https://doi.org/10.1016/j.rser.2018.11.020.
[2] Temoa – tools for energy model optimization and analysis. https://temoacloud.com/. (Accessed 13 January 2021).
[3] OSeMOSYS, OSeMOSYS. http://www.osemosys.org/. (Accessed 13 January 2021).
[4] News | open source initiative. https://opensource.org/. (Accessed 17 February 2021).
[5] Openmod - open energy modelling initiative. https://openmod-initiative.org/manifesto.html. (Accessed 17 February 2021).
[6] M. Heron, V.L. Hanson, I. Ricketts, Open source and accessibility: advantages and limitations, J. Interact. Sci. 1 (1) (May 2013) 2, https://doi.org/10.1186/2194-0827-1-2.
[7] solver:academic_programs_by_solver_partners [GAMS Support Wiki]. https://support.gams.com/solver:academic_programs_by_solver_partners. (Accessed 2 June 2021).
[8] S. R. Thorncraft, H. R. Outhred, and D. J. Clements, Evaluation of Open-Source LP Optimization Codes in Solving Electricity Spot Market Optimization Problems.
[9] B. Meindl, M. Templ, B. Meindl, M. Templ, T. Wien, Analysis of Commercial and Free and Open Source Solvers for Linear Optimization Problems, 2012.
[10] Regional energy deployment system (ReEDS). https://www.nrel.gov/analysis/reeds/index.html. (Accessed 17 January 2021).
[11] Request access to download model. https://www.nrel.gov/analysis/reeds/request-access.html. (Accessed 27 July 2021).
[12] ATB | NREL. https://atb.nrel.gov/. (Accessed 15 October 2021).
[13] A. Martinez, Integrated Canada-U.S. Power Sector Modeling with the Regional Energy Deployment System (ReEDS), National Renewable Energy Laboratory, Golden, Colorado, 2013 [Online]. Available, https://purl.fdlp.gov/GPO/gpo40815. (Accessed 23 April 2021) (["purl.fdlp.gov"]}).
[14] J.L. Ho, W.J. Cole, E. Spyrou, ReEDS-Mexico: A Capacity Expansion Model of the Mexican Power System, National Renewable Energy Lab. (NREL), Golden, CO (United States), Sep. 2017, https://doi.org/10.2172/1398375. NREL/TP-6A20-70076.
[15] A. Rose, Chernyakhovskiy, J. Palchak, S. Koebrich, M. Joshi, Least-Cost Pathways for India's Electric Power Sector, I. (ORCID:0000000284917814), National Renewable Energy Lab. (NREL), Golden, CO (United States), May 2020, https://doi.org/10.2172/1659816. NREL/TP-6A20-76153.
[16] S. Sattler, J. Macknick, D. Yates, F. Flores-Lopez, A. Lopez, J. Rogers, Linking electricity and water models to assess electricity choices at water-relevant scales, Environ. Res. Lett. 7 (4) (Dec. 2012) 45804, https://doi.org/10.1088/1748-9326/7/4/045804.
[17] R.L. Newmark, S.M. Cohen, K. Averyt, J. Macknick, J. Meldrum, P. Sullivan, Evaluating the Impacts of Climate Change on the Operations and Future Development of the U.S. Electricity System, vol. 2014, Dec. 2014. GC22A-01.
[18] E. Zhou, W. Cole, B. Frew, Valuing variable renewable energy for peak demand requirements, Energy 165 (Dec. 2018) 499–511, https://doi.org/10.1016/j.energy.2018.09.009.
[19] E. Ibanez, T. Mai, L. Coles, Analyzing the Deployment of Large Amounts of Offshore Wind to Design an Offshore Transmission Grid in the United States: Preprint, National Renewable Energy Lab. (NREL), Golden, CO (United States), Sep. 2012. NREL/CP-5500-56262, https://www.osti.gov/biblio/1051891. (Accessed 15 October 2021) [Online]. Available:.
[20] O. Zinaman, E. Ibanez, D. Heimiller, K. Eurek, T. Mai, Modeling the Integrated Expansion of the Canadian and U.S. Power Sectors with the Regional Energy Deployment System (ReEDS), National Renewable Energy Lab. (NREL), Golden, CO (United States), Jul. 2015, https://doi.org/10.2172/1215138. NREL/TP-6A20-63797.
[21] Y. Scholz, et al., Speeding up Energy System Models - a Best Practice Guide, Jun. 2020. https://gitlab.com/beam-me/bpg. (Accessed 18 January 2021).

[22] Standard scenarios. https://www.nrel.gov/analysis/standard-scenarios.html. (Accessed 19 April 2021).
[23] E. Klotz, A.M. Newman, Practical guidelines for solving difficult mixed integer linear programs, Surv. Oper. Res. Manag. Sci. 18 (1) (Oct. 2013) 18–32, https://doi.org/10.1016/j.sorms.2012.12.001.
[24] BDMLP. https://new.gams.com/33/docs/S_BDMLP.html. (Accessed 23 April 2021).
[25] coin-or/Clp, COIN-OR Foundation, 2021 [Online]. Available: https://github.com/coin-or/Clp. (Accessed 2 June 2021).
[26] CLP user guide. https://www.coin-or.org/Clp/userguide/clpuserguide.html#presolve. (Accessed 25 March 2021).
[27] GLPK - GNU project - free software foundation (FSF). https://www.gnu.org/software/glpk/. (Accessed 2 June 2021).
[28] Ipopt: documentation. https://coin-or.github.io/Ipopt/. (Accessed 17 January 2021).
[29] Introduction. http://lpsolve.sourceforge.net/5.5/Intro.htm. (Accessed 2 June 2021).
[30] MINOS 5.5. http://www.sbsi-sol-optimize.com/asp/sol_product_minos.htm. (Accessed 2 June 2021).
[31] MINOS - how to buy. http://www.sbsi-sol-optimize.com/asp/sol_minos.htm. (Accessed 21 July 2021).
[32] Argonne-National-Laboratory/PIPS, Argonne National Laboratory, 2021 [Online]. Available: https://github.com/Argonne-National-Laboratory/PIPS. (Accessed 2 June 2021).
[33] BEAM-ME - project. http://www.beam-me-projekt.de/beam-me/EN/Project/project_node.html. (Accessed 23 April 2021).
[34] SoPlex. https://soplex.zib.de/. (Accessed 2 June 2021).
[35] SoPlex: overview. https://soplex.zib.de/doc-2.2.1/html/index.php. (Accessed 21 July 2021).
[36] coin-or/Ipopt, COIN-OR Foundation, 2021 [Online]. Available: https://github.com/coin-or/Ipopt. (Accessed 2 June 2021).
[37] Solver manuals. https://www.gams.com/latest/docs/S_MAIN.html. (Accessed 22 January 2021).
[38] J.L. Gearhart, K.L. Adair, J.D. Durfee, K.A. Jones, N. Martin, R.J. Detry, Comparison of Open-Source Linear Programming Solvers, Sandia National Lab. (SNL-NM), Albuquerque, NM (United States), Oct. 2013, https://doi.org/10.2172/1104761. SAND2013-8847.
[39] Decison tree for optimization software. http://plato.asu.edu/bench.html. (Accessed 17 January 2021).
[40] O. US EPA, Sources of Greenhouse Gas Emissions, *US EPA*, Dec. 29, 2015. https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions. (Accessed 22 January 2021).
[41] D. Getman, A. Lopez, T. Mai, M. Dyson, Methodology for Clustering High-Resolution Spatiotemporal Solar Resource Data, National Renewable Energy Lab. (NREL), Golden, CO (United States), Sep. 2015, https://doi.org/10.2172/1225896. NREL/TP-6A20-63148.
[42] Intel® Xeon® Gold 5120 Processor (19.25M Cache, 2.20 GHz) - Product Specifications, Intel, in: https://www.intel.com/content/www/us/en/products/sku/120474/intel-xeon-gold-5120-processor-19-25m-cache-2-20-ghz/specifications.html. (Accessed 2 May 2021).
[43] B.A. Frew, M.Z. Jacobson, Temporal and spatial tradeoffs in power system modeling with assumptions about storage: an application of the POWER model, Energy 117 (Dec. 2016) 198–213, https://doi.org/10.1016/j.energy.2016.10.074.
[44] T. Mai, C. Barrows, A. Lopez, E. Hale, M. Dyson, K. Eurek, Implications of Model Structure and Detail for Utility Planning. Scenario Case Studies Using the Resource Planning Model, National Renewable Energy Lab. (NREL), Golden, CO (United States), Apr. 2015, https://doi.org/10.2172/1215187. NREL/TP-6A20-63972.