Engineering, **E**conomic, and **E**nvironmental **E**lectricity **S**imulation **T**ool: Model Formulation

E4ST Team

May 2022

Contents

1 Introduction

At the core of the E4ST simulation is a linear program which represents the collective behavior of the system operators, electricity end-users, generators, storage units, and investors. The linear program can be described as follows, with the equation numbers in parentheses referencing the equations in this document which define that aspect of the program:

maximize	consumer benefits minus costs of power system operation and investment	(1)
subject to	electricity demand constraints	(17)
	generator capacity constraints	(12) - (13)
	generator availability constraints	(14) - (16)
	storage capacity constraints	(18) - (19)
	storage operation constraints	(20) - (22), (25)
	transmission line limits	(27) - (29), (31)
	power balance constraints	(30)
	carbon capture and storage limits	(33) - (35)
	state and federal energy policies	(39) - (41)

The following sections provide further detail on the design of a single-year E4ST simulation. Section 2 lists the mathematical notation used in this document to describe the model. Sections 3 and 4 describe the core components of the objective function and constraints which make up the linear program. Section 5 describes the various power sector policies which may be added to an E4ST simulation, and how they alter the objective function and set of constraints.

Sections 2 through 5 are adapted from Mao 2017, an earlier description of the E4ST model. This document expands on that work, adding developments that have happened since its publication.

E4ST is built on top of Matpower, an open-source Matlab package for power system simulation (Zimmerman, C. Murillo-Sánchez, and Thomas 2011). The power flow design in Section 4.4 is derived from Zimmerman and C. E. Murillo-Sánchez 2020.

2 Notation

The sets, indices, parameters and variables used in the model formulation are as follows:

Sets and Indices

$i \in \mathcal{G}$ Set of all generators, including direct air capture units [1]

- $\mathcal{G}_e \subset \mathcal{G}$ Set of all existing generators
- $\mathcal{G}_n \subset \mathcal{G}$ Set of all new and buildable generators
- $\mathcal{G}_j \subset \mathcal{G}$ Set of generators which connect to the grid at node j
- $\mathcal{G}^+ \subset \mathcal{G}$ Set of all generators which produce electricity, which are all generators other than direct air capture units
- $\mathcal{G}^{\mathrm{DAC}} \subset \mathcal{G}$ Set of all direct air capture units

[1] For simplicity, we include direct air capture (DAC) units under the 'generator' umbrella term. This is because DAC units act in the model very similarly to other generators, with main difference being that their operation consumes electricity.

$\mathcal{G}_{\text{refit}} \subset \mathcal{G}_e$ Set of generators which can be retrofitted with CCUS.

$\mathcal{G}_i^{\text{refit}} \subset \mathcal{G}_n$ Set of buildable generators which represent the CCUS options for retrofitting generator i

$b \in \mathcal{B}$ Set of all diurnal electricity storage devices

$\mathcal{B}_j \subset \mathcal{B}$ Set of diurnal storage units which connect to the grid at node j

$j \in \mathcal{N}$ Set of all nodes

$\mathcal{N}_j \subset \mathcal{N}$ The set of nodes that are connected to node j with a transmission line

$t \in \mathcal{T}$ Set of all representative days, where each day is an ordered set of representative hours

$k \in \mathcal{H}$ Set of all representative hours

$\{t, n\} \in \mathcal{H}_t$ Set of representative hours in representative day t. Where hour $\{t, n\}$ is the n-th hour in day t.

$\{j, j'\} \in \mathcal{L}$ Set of all transmission lines, where line $\{j, j'\}$ is the line between nodes j and j'

$\delta \in \Delta$ Set of all direct-current (DC) transmission lines

$z \in \mathcal{Z}$ Set of CCUS end-use options

$m \in \mathcal{M}_{pol}$ Set of all policy capacity constraints

$m \in \mathcal{M}_{ccus}$ Set of CCUS capacity constraints

$\psi \in \Psi$ Set of all multi-hour capacity factor constraints

$\upsilon \in \Upsilon$ Set of all interfaces with flow constraints

$\pi \in \Pi$ Set of all emissions constraints

$\phi \in \Phi$ Set of all RPS/CES constraints

Decision Variables

g_{ik} Generator operation for generator i and hour k. For DAC units, this represents electricity consumption. For all other generators, this represents electricity production. [2]

g_{bk} Electricity released from diurnal storage unit b in hour k (MW or MWh)

s_{bk} Electricity stored into diurnal storage unit b in hour k (MW or MWh)

$s_{b\{t,0\}}$ Initial stored energy in diurnal storage unit b and representative day t

d_{jk} Electricity demands at bus node j and hour k (MW or MWh)

r_i Amount of capacity retirement for generator i (MW)

r_b Amount of capacity retirement for diurnal storage unit b (MW)

a_i Amount of capacity addition for generator i (MW)

a_b Amount of capacity addition for diurnal storage unit b (MW)

Θ_{jk} Phase angle of voltage at node j in hour k (degrees or radians)

$f_{\delta k}$ Power flow through DC transmission line δ in hour k (MW)

[2] Because E4ST uses representative hours, a generator's output during each representative hour can be interpreted as power (MW) during the hour or energy (MWh) produced during the hour.

$\mathbf{G}$ Vector variable of generator operation for all generators i and electricity release for all storage units b, for all hours k

$\mathbf{D}$ Vector variable of demands for all nodes j and hours k

$\mathbf{S}$ Vector variable of electricity stored for all storage units b and all hours k

$\mathbf{R}$ Vector variable of capacity retirement for all generators i and storage units b

$\mathbf{A}$ Vector variable of capacity addition for all generators i and storage units b

$\boldsymbol{\Theta}$ Vector variable of phase angle of voltage for all nodes j and hours k

$\boldsymbol{\Delta}$ Vector variable of power flow for all DC transmission lines δ and hours k

Dual Variables

λ_{jk} Lagrange multiplier on the power flow balance constraint at node j in hour k

λ_{π} Lagrange multiplier on emissions constraint π

λ_{ϕ} Lagrange multiplier on RPS/CES constraint ϕ

λ_{m} Lagrange multiplier on capacity policy constraint m

Parameters

p_i^0 Capacity of generator i existing before the current simulation (MW)

c_i^V Non-fuel variable operating and maintenance cost for generator i (\$/MWh)

c_i^{FL} Fuel cost for generator i (\$/MWh)

f_i Fuel price for generator i (\$/MMbtu)

Q_i Heat rate (efficiency) of generator i (MMbtu/MWh)

D_{jk}^{+} Target upper bound of electricity demand at node j in hour k

ω_k Weight of representative hour k

h_k Number of hours represented by hour k

N_t Number of representative hours in representative day t

$h_{\{t,n\}}^{daily}$ The amount of real daily hours of day t represented by hour $\{t, n\}$.

T_b Time it take for diurnal storage unit b to fully discharge at its maximum capacity (hours)

$e_i^{\mathfrak{p}}$ Emission rate of pollutant $\mathfrak{p}$ for generator i (short ton/MWh)

${\rho_i}^{\mathfrak{p}}$ Emission tax for pollutant $\mathfrak{p}$ for generator i (\$/short ton)

e_i^{ccus} CO_2 available for carbon utilization and storage as an output of generator i (short ton/MWh)

c_i^{ccus} Total transportation and end-use cost for captured CO_2 (\$/short ton)

${\rho_i}^{ptc}$ Production tax credit from government to generator i (\$/MWh)

ρ_i^{itc} Investment tax credit from government to generator i, as a percentage of investment costs (%)

c_i^F Fixed operating and maintenance costs for generator i (\$/MW/hr)

c_i^I Investment cost for generator i (\$/MW/hr)

c_i^{LE}	Lifetime extension cost for generator i (\$/MW/hr).
γ_{ik}	Availability factor for generator i in hour k (%)
α_i^{min}	Minimum electricity output percentage for generator i (%)
a_i^+	Maximum capacity that can be built for generator i (MW)
ξ_b	Round-trip storage efficiency of diurnal storage unit b (%)
A_m^-	Lower limit of capacity constraint m (MW)
A_m^+	Upper limit of capacity constraint m (MW)
K_ψ	Maximum capacity factor for constraint ψ (%)
$S_{\{j,j'\}}$	Susceptance of the line between node j to node j' (?)
$F_{\{j,j'\}}$	Flow limit of the transmission line between node j and node j' (MW)
$F_\delta^{\min}$	Lower bound of power flow over DC transmission line δ (MW)
$F_\delta^{\max}$	Upper bound of power flow over DC transmission line δ (MW)
from_δ	The "From" node for DC transmission line δ
to_δ	The "To" node for DC transmission line δ
$F_\upsilon^{\min}$	Lower bound of interface power flow for interface υ (MW)
$F_\upsilon^{\max}$	Upper bound of interface power flow for interface υ (MW)
L_z	Annual CO_2 storage capacity limit for CCUS end-use option z (short tons/year)
ϖ_i	Capacity penalty of retrofitting generator i with CCUS (%)
I_i^z	Indicator parameter for if generator i sends its captured CO_2 to CCUS option z.
I_i^m	Indicator parameter for if generator i is included in capacity constraint m
I_{ik}^π	Indicator parameter for if emissions from generator i and hour k are included in emissions cap π
E_π	Maximum emissions for emissions cap π
φ_i^ϕ	The rate at which generator i earns credits towards RPS/CES policy ϕ
Ω_j^ϕ	The percentage of electricity demand at node j which must be covered by credits under RPS/CES policy ϕ
I_{ik}^ψ	Indicator parameter for if generator i in hour k is included in capacity factor constraint ψ

Functions

$F(\cdot)$	Total gross consumer benefits function for all demands j and hours k
$B_{jk}(\cdot)$	Consumer benefit function for demand at node j and hour k
$C(\cdot)$	Total cost function for all units (generators, DAC, and diurnal storage units) and all hours
$C_{VOM}(\cdot)$	Total variable operating and maintenance cost for all units and hours
$C_{FL}(\cdot)$	Total fuel cost for all units and hours

$C_{FOM}(\cdot)$ Total fixed operating and maintenance cost for all units

$C_{INV}(\cdot)$ Total investment cost for all units

$C_{LE}(\cdot)$ Total lifetime extension costs for all units

$C_{CCUS}(\cdot)$ Total transportation and end-use cost of captured CO_2 for all CCUS units

$C_{POL}(\cdot)$ Total policy-related cost for all units and hours

$C_{POL}^{\mathfrak{p}}(\cdot)$ Total tax on emission $\mathfrak{p}$ for all generators i and hours k

$C_{POL}^{ptc}(\cdot)$ Total production tax credit for all generators and hours k

$C_{POL}^{itc}(\cdot)$ Total investment tax credit for all generators

$E_{CO2}(\cdot)$ Total CO2 emissions for all generators i and hours k

$f_{\{j,j'\}}^{k}(\cdot)$ Power flow from node j to node j' in hour k

$F_{\upsilon}^{k}(\cdot)$ Total directed power flow across interface υ in hour k

Derived Results

LMP_{jk} Locational marginal electricity price at node j and hour k.

2.1 Temporal Representation and Units

A single instance of E4ST is designed to simulate a single year. However, because of the detailed spatial resolution of E4ST, modeling operation in all 8760 hours of the year would make the model too large to solve practically and quickly. We instead model operation over a set of several representative days, each of which consists of a set of representative hours. The representative days and hours are chosen to represent closely the frequency distributions of load, wind availability, and solar availability throughout the time period, as well as capture the periods of highest generation scarcity. Generation scarcity can come from periods of high load, low wind or solar availability, and combinations of these conditions.

The set of days is denoted by $\mathcal{T}$ and indexed by t, and the set of hours is denoted by $\mathcal{H}$ and indexed by k. To closely match a real distribution of load, wind, and solar, each representative hour is weighted so that the set of hours itself is a probability distribution. The weight of each hour $k \in \mathcal{H}$ is denoted by ω_k, where $\sum_{k\in\mathcal{H}} \omega_k = 1$. We can also think about the weight of each hour in terms of how many hours of the year it represents, denoted by h_k, where $h_k = 8760 * \omega_k \quad \forall k \in \mathcal{H}$.

We use ω_k, rather than h_k, when formulating E4ST's objective function. This means that the components of E4ST's objective function value can be interpreted as the benefits and costs of a single hour of operation, where that hour is the average of the hours in the year. This helps keep down the magnitude of the objective function value. To calculate the total benefits and costs for the year, we can multiply those values by 8760.

For consistency, capacity-related decision variables are in terms of MW, and production-related decision variables are in terms of MWh. Thus, variable costs that scale with generation use units of dollars per MWh ($/MWh), while fixed costs that scale with capacity use units of dollars per MW per hour ($/MW/hr), because the objective function is benefits and costs of an average hour, not of the full year. This allows us to have all units of cost comparable on the level of a single hour. Note that the units ($/MWh) and ($/MW/hr) appear similar but are not the same.

3 Objective Function

The model maximizes total gross consumer benefits minus the total cost, as shown in the following forms:

$$\max_{\mathbf{G},\mathbf{D},\mathbf{S},\mathbf{R},\mathbf{A},\mathbf{\Theta},\mathbf{\Delta}} F(\mathbf{D}) - C(\mathbf{G},\mathbf{R},\mathbf{A}) \tag{1}$$

where $F(\mathbf{D})$ is the total gross consumer benefits function for all demands j and hours k; $C(\mathbf{G}, \mathbf{R}, \mathbf{A})$ is the total cost function for all power system components in all hours k, including all generators i and diurnal storage units b.

The optimization variables are as follows:

$$\mathbf{x} = \begin{bmatrix} \mathbf{G} \\ \mathbf{D} \\ \mathbf{S} \\ \mathbf{R} \\ \mathbf{A} \\ \mathbf{\Theta} \\ \mathbf{\Delta} \end{bmatrix} \tag{2}$$

where $\mathbf{G}$ is the vector variable of generator outputs for all generators i, and electricity release for all storage units b, for all hours k; $\mathbf{D}$ is the vector variable of demands for all nodes j and hours k; $\mathbf{S}$ is the vector variable of electricity stored for all storage units b and hours k; $\mathbf{R}$ is the vector variable of capacity retirement for all generators i; $\mathbf{A}$ is the vector variable of capacity addition for all generators i; $\mathbf{\Theta}$ is the vector variable of phase angle of voltage for all nodes j and hours k; $\mathbf{\Delta}$ is the vector variable of power flow for all DC transmission lines δ and hours k.

Note that not all decision variable appear directly in the objective function (e.g. $\mathbf{S}, \mathbf{\Theta}, \mathbf{\Delta}$). Flow through power lines and into energy storage do not directly reduce system costs unless they cause existing power plants to either (1) produce less electricity or (2) require less capacity.

Note that the revenue gained from operation of the individual units is not explicitly included in the objective function. This revenue would mainly be from selling electricity for generators, selling carbon policy credits for DAC units, or from arbitraging electricity for storage units. This is because under perfect competition, all producers are price takers and therefore the profit maximization problem is identical to cost minimization. In reality, electricity markets have imperfect competition. Nonetheless the system is usually represented using a social planner model like E4ST.

In the following detailed descriptions of each part of the objective function, the default is that values are summed over the full sets of generators, storage units, nodes, and hours, which means all $i \in \mathcal{G}$, $b \in \mathcal{B}$, $j \in \mathcal{N}$ and $k \in \mathcal{H}$ if there is no other notation in the equation.

3.1 Consumer Benefits

Consumer benefits can be defined as a function of electricity demand d_{jk} at node j for hour k. This function can be a linear or piecewise linear function. Then the total consumer benefit function can be calculated as follows:

$$F(\mathbf{D}) = \sum_{j} \sum_{k} \omega_k B_{jk}(d_{jk}) \tag{3}$$

where $B_{jk}(d_{jk})$ is the consumer benefit function for demand at node j and hour k; ω_k is the weight of hour k.

What is meant by a benefit function? This can be demonstrated by a simple example. Suppose that each consumer is willing to consume electricity normally if the electricity price is less than 5000 \$ per MWh and decides to stop consuming electricity if the electricity price is above this threshold. While this may sound unrealistic, it corresponds to a scenario in which short-term electricity demand is inelastic (e.g. because consumers are not exposed to real time prices), yet utilities shed load if the electricity price exceeds a threshold value. This makes the value 5000 \$ per MWh the marginal benefit of consumption since it is the most consumers are willing to pay, and at higher prices, consumers will no longer find it beneficial to consume electricity. In this case the benefit function would simply be:

$$B_{jk} = 5000 \cdot d_{jk} \tag{4}$$

3.2 Power System Costs

The total power-system cost function can be further broken down into variable operating and maintenance cost, fixed operating and maintenance cost, investment and lifetime extension cost, and policy-related cost, as seen in Equation (5). The costs for generators, direct air capture units, and diurnal storage units all follow this general structure, and so can be discussed together. We will use the general term 'units' when we need to refer to something that can be a generator, DAC unit, or diurnal storage unit.

$$
\begin{aligned}
C(\mathbf{G},\mathbf{R},\mathbf{A}) = C_{VOM}(\mathbf{G}) + C_{FL}(\mathbf{G}) + C_{FOM}(\mathbf{R},\mathbf{A}) \\
+ C_{INV}(\mathbf{A}) + C_{LE}(\mathbf{R}) + C_{CCUS}(\mathbf{G}) + C_{POL}(\mathbf{G},\mathbf{A})
\end{aligned}
\tag{5}
$$

where $C_{VOM}(\mathbf{G})$ is the total variable operating and maintenance cost for all units and hours; $C_{FL}(\mathbf{G})$ is the total fuel cost for all units and hours; $C_{FOM}(\mathbf{R},\mathbf{A})$ is the total fixed operating and maintenance cost for all units; $C_{INV}(\mathbf{A})$ is the total investment cost for all units; $C_{LE}(\mathbf{R})$ is the total lifetime extension costs for all units; $C_{CCUS}(\mathbf{G})$ is the total CCUS transportation and end-use cost for all units; and $C_{POL}(\mathbf{G},\mathbf{A})$ is the total policy-related cost for all units, hours, and active policies.

We use $C_{POL}(\mathbf{G},\mathbf{A})$ as a placeholder to refer to all the ways that policies may impact the objective function. In section 5, we list the policies that can be added onto an instance of E4ST, and how they affect the objective function.

Also, note that $C_{CCUS}(\mathbf{G})$ is described in Section 4.5 instead of Section 3.2.

> All cost parameters defined in Section 3.2 for generators, indexed by generator i, are also defined correspondingly for diurnal storage units, indexed by storage unit b. The exception is that diurnal storage units do not have heat rate or fuel cost parameters. DAC units may have heat rate and fuel cost parameters as some DAC types require fuel input.

3.2.1 Variable Operating and Maintenance Cost

Variable operating and maintenance (VOM) cost refers to all non-fuel costs which scale depending on the amount generated, and can be defined as a function of generator output g_{ik} for generator i and hour k. Similarly, for DAC units, VOM costs scale with the electricity consumed. For diurnal storage units, we choose to specify VOM costs in terms of dollars per MWh discharged, so VOM costs for diurnal storage also scale with g_{bk}.

E4ST uses a simple linear function as follows:

$$
C_{VOM}(\mathbf{G}) = \sum_i \sum_k \omega_k c_i^V g_{ik} + \sum_b \sum_k \omega_k c_b^V g_{bk} \tag{6}
$$

where c_i^V is the non-fuel variable cost per megawatt-hour for generator i, and c_b^V is the variable cost per megawatt-hour discharged for storage unit b.

3.2.2 Fuel Cost

Fuel cost can also be defined as a function of generator output g_{ik} for generator i and hour k. Only some generators will have a non-zero fuel cost, others, particularly reneable resources, tend to require no fuel to operate. DAC units typically do not require fuel, but some types may. Diurnal storage does not require fuel, and so has no fuel cost.

E4ST uses a simple linear function as follows:

$$
C_{FL}(\mathbf{G}) = \sum_i \sum_k \omega_k c_i^{FL} g_{ik} \tag{7}
$$

where c_i^{FL} is the fuel cost per megawatt-hour for generator i.

Actual data on fuel costs is usually give in terms of \$/MMbtu. MMbtu is equivalent to one million british thermal units of energy. It is used as a measure of *heat input*, meaning that it measures total energy released during the combustion of a fuel. The actual fuel cost per MWh for a generator therefore depends not only on the raw fuel price but also the the efficiency of the generator. We calculate the fuel cost per megawatt-hour for generator i as

$$c_i^{FL} = f_i * Q_i \tag{8}$$

where f_i is the fuel price for generator i, in \$/MMbtu, and Q_i is the heat rate of generator i, in MMbtu/MWh. The heat rate is a measure of efficiency since it is the ratio of heat input to energy output.

3.2.3 Fixed Operating and Maintenance Cost

Fixed operating and maintenance (FOM) cost are the costs that scale with a unit's capacity, regardless of how much the capacity is being used. FOM costs can be defined as a function of the actual active capacity for generator i, which can be determined by capacity retirement r_i and capacity addition a_i.

$$C_{FOM}(\mathbf{R}, \mathbf{A}) = \sum_i c_i^F (p_i^0 - r_i + a_i) + \sum_b c_b^F (p_b^0 - r_b + a_b) \tag{9}$$

where p_i^0 is the initial capacity for generator i. Note that $p_i^0 = 0$ if generator i is a newly built generator; capacity retirement $r_i = 0$ if generator i is a newly built generator; capacity addition $a_i = 0$ if generator i is an existing generator; c_i^F is the fixed operating and maintenance costs per megawatt per hour for generator i.

3.2.4 Investment Cost

Investment cost can be defined as a function of capacity addition a_i. Total investment cost (in \$/MW) is amortized over an assumed economic lifetime and then it is converted to an hourly measure of investment cost in \$/MW/hr. The total investment cost is added to the objective function as follows:

$$C_{INV}(\mathbf{A}) = \sum_i c_i^I a_i + \sum_b c_b^I a_b \tag{10}$$

where c_i^I is the investment cost per megawatt per hour for generator i.

3.2.5 Lifetime Extension Cost

Pre-existing generators often have to make investments once they reach a certain age in order to keep operating. These could be investments in replacing equipment, as well as cost for recertification to keep operating. Like investment costs, these costs are amortized, so they can be paid over several years. These costs are a function of capacity which is kept online in the simulation. Diurnal storage units typically do not have life extension costs, but they can be applied in the same way if necessary. Total lifetime extension cost can be calculated as follows:

$$C_{LE}(\mathbf{R}) = \sum_i c_i^{LE} (p_i^0 - r_i) \tag{11}$$

where c_i^{LE} is the lifetime extension cost per megawatt per hour for generator i.

4 Constraints

4.1 Generator Constraints

4.1.1 Generator Capacity Constraint

For all generators $i \in \mathcal{G}$, the quantity of capacity retired from generator i cannot be larger than its initial capacity p_i^0 at the beginning of the simulation. Also, a generator cannot retire a negative amount of capacity. This leads to the follow constraint:

$$0 \le r_i \le p_i^0 \quad \forall i \in \mathcal{G} \tag{12}$$

For new generators that did not exist at the beginning of a simulation, $p_i^0 = 0$.

Similarly, each generator i has an individual maximum capacity limit a_i^+ on the new capacity that can be added. Also, a generator cannot add a negative amount of capacity. This leads to the following constraint:

$$0 \le a_i \le a_i^+ \quad \forall i \in \mathcal{G} \tag{13}$$

For simplicity in E4ST, we only allow capacity additions at newly constructed generators, so $a_i^+ = 0 \quad \forall i \in \mathcal{G}_e$, where $\mathcal{G}_e$ represents the set of pre-existing generators.

4.1.2 Availability Factor Constraint

Generators cannot always operate at their nameplate capacity. The generation from itermittent renewables such as wind and solar depends on time of day and weather conditions, and firm generation such as nuclear and fossil fuels need to be down for a certain percentage of the time for planned maintenance. So, there are upper bounds for electricity outputs from generator i in each hour k, which equal the total active capacity of i multiplied by its availability factor in hour k. The upper bound generator output constraint is shown in Equation (14).

For some types of generators such as coal and hydro, there are lower bounds to ensure that minimum electricity outputs of the generator i in hour k are no less than a fraction of the dispatched capacity. The lower bound generator output constraint is shown in Equation (15).

$$g_{ik} \le \gamma_{ik}(p_i^0 - r_i + a_i) \quad \forall i \in \mathcal{G}, k \in \mathcal{H} \tag{14}$$

$$g_{ik} \ge \alpha_i^{min}(p_i^0 - r_i + a_i) \quad \forall i \in \mathcal{G}, k \in \mathcal{H} \tag{15}$$

where γ_{ik} is the availability factor for generator i in hour k; α_i^{min} is the minimum electricity output percentage for generator i.

4.1.3 Capacity Factor Constraint

Capacity factor is the ratio of the actual electricity output to the maximum possible output for a generator over the course of a year. The capacity factor of a generator differs from its availability factor in that the capacity factor is a long-term average that depends on operation while availability factors are a property of each individual generator and hour. The user may want to set limits on a generator's individual or a set of generators' aggregate seasonal or yearly capacity factor. This is particularly useful for modeling the operation of hydropower generators, where several generators in an area share the same pool of fuel (water), and whose realistic outputs may depend on the season or be well below the maximum possible output. Optionally, the user can give each generator or group more than one capacity factor constraint, such as one for each part of the year, and these constraints can overlap:

$$\sum_i \sum_k \omega_k I_{ik}^{\psi} g_{ik} \le K_{\psi} \sum_i \sum_k \omega_k I_{ik}^{\psi}(p_i^0 - r_i + a_i) \quad \forall \psi \in \Psi \tag{16}$$

where I_{ik}^{ψ} is an indicator parameter which equals 1 if generator i in hour k is included in the constraint ψ, and equals 0 otherwise. K_{ψ} is the corresponding maximum capacity factor for constraint ψ.

4.2 Demand Constraints

In each E4ST simulation, there is a target amount of electricity demand, or load, D_{jk}^+ at each node j and each representative hour k. We assume that the quantity of electricity demanded at each node and hour, d_{jk}, can be reduced (curtailed) if that would be beneficial to the objective function given the cost of supplying generation and the benefit of electricity demanded. However, we do not allow demands to exceed their target amount. A node also cannot demand an amount of electricity less than zero. At that point, it would become a generator. This is enforced through the following constraint.

$$0 \le d_{jk} \le D_{jk}^+ \quad \forall j \in \mathcal{N}, k \in \mathcal{H} \tag{17}$$

4.3 Diurnal Storage Constraints

For capacity added and retired, diurnal storage units follow the same constraints defined in section 4.1.1 for generators, adapted below:

$$0 \leq r_b \leq p_b^0 \quad \forall b \in \mathcal{B} \tag{18}$$

$$0 \leq a_b \leq a_b^+ \quad \forall b \in \mathcal{B} \tag{19}$$

However, the operation of diurnal storage units requires a set of constraints to represent the storage, discharge, and loss of electricity over time. First, at any given time, the storage unit cannot charge or discharge at a rate faster than its rated capacity, enforced by the following constraints:

$$0 \leq g_{bk} \leq (p_b^0 - r_b + a_b) \quad \forall b \in \mathcal{B}, k \in \mathcal{H} \tag{20}$$

$$0 \leq s_{bk} \leq (p_b^0 - r_b + a_b) \quad \forall b \in \mathcal{B}, k \in \mathcal{H} \tag{21}$$

We model diurnal storage, which completes its charging and discharging cycles over timeframes on the scale of a single day, rather than longer-term storage, which discharge energy which was charged days, weeks, or even seasons before. So, we require that over the course of each day, the diurnal storage unit must discharge as much energy as it charges, minus the round-trip storage efficiency of the unit. The round-trip efficiency of unit b, ξ_b, is defined as a percentage of how much of the charged energy is available for discharging after loss. Instead of using k to index the hours, we will use the ordered set $\{t, n\}$ to denote the n-th hour in representative day t. N_t is the number of representative hours in day t.

$$\sum_{n=1}^{N_t} \omega_{\{t,n\}}(\xi_b\, s_{b\{t,n\}} - \, g_{b\{t,n\}}) = 0 \quad \forall b \in \mathcal{B}, t \in \mathcal{T} \tag{22}$$

From constraint (22) it can be shown that the total lost energy in a year due to the round trip efficiency of storage unit b can be written as:

$$\sum_k h_k(s_{bk} - g_{bk}) = (1 - \xi_b) \sum_k h_k s_{bk} \tag{23}$$

the simpler formulation of the loss, on the right hand side of (23) will be useful later.

Finally, the total energy content of the battery must always remain between zero and its energy storage limit. A storage unit's energy storage limit can be calculated based on the unit's duration T_b, which is the number of hours it would take for storage unit b to fully discharge at its maximum capacity. In other words, T_b measures the MWh of energy that is able to be discharged per MW of the unit's capacity. The energy storage limit is then calculated by $T_b(p_b^0 - r_b + a_b)$. Also, because there are typically less than 24 representative hours in a representative day, we need to know how many real hours in a day each representative hour would represent, so that we know how much electricity is added or removed from the storage unit's energy content due to operation in each representative hour. The real daily hours can be calculated from the hour weights as follows:

$$h_{\{t,n\}}^{daily} = 24 * \frac{\omega_{\{t,n\}}}{\sum_{n=1}^{N_t} \omega_{\{t,n\}}} \tag{24}$$

We constrain the energy stored in the unit at the end of each hour to lie within the energy storage limits using the following set of constraints:

$$0 \leq s_{b\{t,0\}} - \sum_{n=1}^{\hat{n}} h_{\{t,n\}}^{daily}(g_{b\{t,n\}} - \xi_b\, s_{b\{t,n\}}) \leq T_b(p_b^0 - r_b + a_b)$$
$$\forall b \in \mathcal{B}, t \in \mathcal{T}, \hat{n} \in \{1, 2, \ldots, N_t\} \tag{25}$$

The initial stored energy $s_{b\{t,0\}}$ for unit b and day t is a decision variable that is implictly constrained by constraints (25) to lie within the energy limits of the unit.

4.4 Power Flow and Nodal Balance Constraints

4.4.1 AC Lines

The electricity transmission system in E4ST consists of a set of transmission lines $\mathcal{L}$, where each line is referred to by an unordered set $\{j, j'\} \in \mathcal{L}$, where line $\{j, j'\}$ is the transmission line between the two nodes j and j'. Each transmission line has a susceptance $S_{\{j,j'\}}$. Because $\{j, j'\}$ and $\{j', j\}$ refer to the same transmission line between the nodes, $S_{\{j,j'\}} = S_{\{j',j\}}$.

The power flow $f^k_{\{j,j'\}}(\Theta_{jk}, \Theta_{j'k})$ from node j to node j' in hour k is modeled as the susceptance of the line between node j and node j', $S_{\{j,j'\}}$, multiplied by the phase angle Θ_{jk} difference between these two nodes:

$$f^k_{\{j,j'\}}(\Theta_{jk}, \Theta_{j'k}) = S_{\{j,j'\}}(\Theta_{jk} - \Theta_{j'k}) \tag{26}$$

From Equation (26), the phase angle difference drives power flow between two nodes. Even though each transmission line $\{j, j'\}$ is not directed, and usually can flow in either direction, the measurement of flow along that line does have direction, so order does matter in $f^k_{\{j,j'\}}$. Note that $f^k_{\{j,j'\}} = -f^k_{\{j',j\}}$.

The electricity flow on the line between nodes j and j' cannot exceed the flow limit of the line, $F_{\{j,j'\}}$.

$$-F_{\{j,j'\}} \leq S_{\{j,j'\}}(\Theta_{jk} - \Theta_{j'k}) \leq F_{\{j,j'\}} \quad \forall \{j, j'\} \in \mathcal{L}, k \in \mathcal{H} \tag{27}$$

One node in the system must serve as a reference node, with a fixed phase angle, so that the system has a unique solution. Denoting the reference node as node $\hat{j}$. We add the following constraint:

$$\Theta_{\hat{j}} = 0 \tag{28}$$

At all other buses other than the reference bus, phase angles are unrestricted, and have no bounds.

4.4.2 DC Lines

In E4ST, we represent direct current transmission lines as separable from the operation of the AC transmission grid. A DC line in E4ST can transfer power between two nodes, as long that amount of power is between its power flow limits, which are pre-determined in an E4ST simulation. We represent this by including the DC line power flows in the nodal balance constraints, and implementing the following constraint:

$$F^{\min}_{\delta} \leq f_{\delta k} \leq F^{\max}_{\delta} \quad \forall \delta \in \Delta, k \in \mathcal{H} \tag{29}$$

where $F^{\min}_{\delta}$ and $F^{\max}_{\delta}$ are the upper and lower bounds of power flow across the DC line. Each DC line in E4ST is directed. DC line δ goes from its from node from_{δ} to its to node to_{δ}. This directionality is necessary to define the flow, but a single DC line can flow both ways, by allowing the lower bound on its power flow to be negative.

4.4.3 Nodal Balance

At every node of the electricity system, all electricity should be accounted for in each hour. The net electricity generation and use at each node (including output from generators, use by DAC units, charging and discharging from storage units, and general electricity demand at the node) should equal the net power flowing out of the node, which is shown in an equality constraint as follows:

$$\sum_{i \in \mathcal{G}^+_j} g_{ik} - \sum_{i \in \mathcal{G}^{DAC}_j} g_{ik} + \sum_{b \in \mathcal{B}_j} (g_{bk} - s_{bk}) - d_{jk} - \sum_{j' \in \mathcal{N}_j} S_{\{j,j'\}}(\Theta_{jk} - \Theta_{j'k}) + \sum_{\delta:\text{to}_{\delta}=j} f_{\delta k} - \sum_{\delta:\text{from}_{\delta}=j} f_{\delta k} = 0$$
$$\forall j \in \mathcal{N}, k \in \mathcal{H} \tag{30}$$

where $\mathcal{G}^+_j$ is the set of electricity-producing generators that connect to the grid at node j; $\mathcal{G}^{DAC}_j$ is the set of direct air capture devices that connect to the grid at node j; $\mathcal{B}_j$ is the set of diurnal storage units that connect to the grid at node j; $\mathcal{N}_j$ is the set of nodes connected to node j with a transmission line.

4.4.4 Interface Flow Limits

Another power flow constraint is the interface flow limit. An interface flow is the sum of directed power flows in a set of transmission lines. The power flow in a single transmission line in hour k can be calculated based on Equation (26). Then the flow interface constraints can be defined as follows (Zimmerman & Murillo-Sánchez, 2016).

$$F_{\upsilon}^{\min} \leq F_{\upsilon}^{k}(\mathbf{\Theta}) \leq F_{\upsilon}^{\max} \quad \forall \upsilon \in \Upsilon, k \in \mathcal{H} \tag{31}$$

where Υ is the set of all interfaces with flow constraints; $F_{\upsilon}^{k}(\mathbf{\Theta})$ is the sum of directed power flow for interface υ in hour k; $F_{\upsilon}^{\min}$ and $F_{\upsilon}^{\max}$ are the lower bound and upper bound on the power flow for interface υ.

4.5 CCUS Constraints

Several types of units in E4ST can participate in carbon capture, utilization, and storage (CCUS). Both fossil-fuel generators equipped with carbon capture equipment and direct air capture units have as an output an amount of captured CO_2 which must be transported, utilized, or sequestered. For these units, we define a new 'emissions' rate, e_i^{ccus}, which is the amount of CO_2, in short tons per MWh of operation, which is available for utilization or storage. Note that e_i^{ccus} is not just the negative of $e_i^{CO_2}$, and $e_i^{CO_2}$ does not have to equal zero for units that participate in CCUS, because generators may still emit some CO_2 while sequestering, and even DAC units may not be perfectly efficient and remove the same amount of net CO_2 from the atmosphere as it is able to deliver as captured CO_2.

The cost of transportation, utilization, and storage of CO_2 can vary depending on start and end location of the CO_2, as well as the end use of the CO_2. Even for a single generator, there may be several options for the end use of the captured CO_2. In E4ST, since generators cannot have two different costs, we model each CCUS option for a generator as an individual generator with its own CCUS cost.

For the costs, we define a cost term c_i^{ccus} for each generator i that particpates in CCUS. This is the total transportation and end use cost per short ton of CO_2 captured. The total cost to be included in the objective function is:

$$C_{CCUS}(\mathbf{G}) = \sum_i \sum_k \omega_k c_i^{ccus} e_i^{ccus} g_{ik} \tag{32}$$

We can define the set of CCUS end-use options as $\mathcal{Z}$, indexed by z. Each option z has an annual limit, L_z on the amount of CO_2 that can be stored there, from all power-sector sources. We implement this limit with the following constraint:

$$\sum_i \sum_k \omega_k I_i^z e_i^{ccus} g_{ik} \leq \frac{L_z}{8760} \quad \forall z \in \mathcal{Z} \tag{33}$$

where I_i^z is an indicator parameter that equals 1 if generator i sends its CO_2 to CCUS option z, and equals 0 otherwise. Note that the yearly capacity is divided by 8760 to get an hourly average limit, which we use because the constraint is formulated using ω_k.

For data on the actual CCUS options available, their costs, capacities, and transportation costs, we borrow the assumptions in Chapter 6 of the EPA's Power Sector Modeling Platform v6 (cite here).

Because we must model a separate generator for each CCUS option, each with their own buildable capacity limit, we may want in some cases to limit the total capacity built at a certain set of CCUS generators. We can do this with the following capacity constraint:

$$\sum_i I_i^m a_i \leq A_m^+ \quad \forall m \in \mathcal{M}_{ccus} \tag{34}$$

where I_i^m is an indicator parameter which equals 1 if generator i is included in capacity constraint m, A_m^+ is the upper limit on the capacity constraint, and $\mathcal{M}_{ccus} \subset \mathcal{M}$ is the set of CCUS capacity constraints.

A variant of this constraint can be used to model CCUS retrofits at existing power plants. We generally allow existing coal power plants the option to be retrofitted for CCUS. In this case, we create a set of buildable CCUS generators, and constrain the sum of their total added capacity and the remaining un-retrofitted capacity to not exceed the original capacity of the existing unit, minus the capacity penalty the generator would experience when retrofitting.

$$(p_i^0 - r_i + a_i) + \frac{1}{\varpi_i} \sum_{m \in \mathcal{G}_i^{\text{refit}}} a_m \leq p_i^0 \quad \forall i \in \mathcal{G}_{\text{refit}} \tag{35}$$

where ϖ_i is the capacity penalty of retrofitting generator i. The capacity of generator i, if it was fully retrofitted, would be ϖ_i multiplied by the original capacity. $\mathcal{G}_i^{\text{refit}}$ is the set of buildable generators which represent the CCUS options for retrofitting generator i, and $\mathcal{G}_{\text{refit}}$ is the set of generators which can be retrofitted. Note that this allows a generator the option to only partially retrofit with CCUS. Also, while the VOM, FOM, and fuel cost rates of the buildable retrofits would be those of the full power plant, after any alterations for the operation of the CCUS, the investment cost of the retrofit units would cover only the capital costs of the retrofit, and not the full capital costs of the entire generator. The emissions rates of the retrofit units would of course be changed to reflect the carbon capture.

5 Policy Additions

5.1 Policies modeled in the Objective Function

We use our objective function to optimize the cost to system operators and the benefit to consumers Policies such as emissions taxes or tax credits should be factored into the objective function if they influence the operating costs. In E4ST, the most common types of policies which affect the objective function are emissions taxes, production tax credits, and investment tax credits.

The descriptions in this section are intentionally general, because any number of policies can be included in an E4ST simulation, potentially applying to different regions, generation types, pollutants, etc. For example, an E4ST simulation may include federal production tax credits on all wind turbines in the US, a tax on CO_2e emissions in California, a tax on NO_x emissions which only applies to generators larger than 100 MW in New England, and no investment tax credits at all.

5.1.1 Emission Tax

In this section, we assume that the emissions rate per MWh for generator i is already determined. Emissions taxes are most commonly applied and proposed for CO_2, but we can apply an emissions tax to any of the emissions tracked by E4ST, which are CO_2, CO_2e, NO_x, SO_2, and $PM_{2.5}$, so we will discuss an emissions tax generally here.

Assuming the emission tax on pollutant $\mathfrak{p}$ is $\rho_i^{\mathfrak{p}}$ for generator i, then the total emission tax is as follows:

$$C_{POL}^{\mathfrak{p}}(\mathbf{G}) = \sum_i \sum_k \omega_k \rho_i^{\mathfrak{p}} e_i^{\mathfrak{p}} g_{ik} \tag{36}$$

If these emissions taxes exist, they are added to the generator cost function $C(\cdot)$ because they increase the percieved cost of operating the generators. This means they are subtracted from the overall objective function.

5.1.2 Production Tax Credit

A production tax credit effectively transfers money from the government to electricity producers for every unit of energy (e.g. MWh) produced. A production subsidy has same effect. These policies reduce the perceived variable cost of generation to the producers. They thus constitute a positive addition in the objective function (a subtraction from the generator cost function):

$$C_{POL}^{ptc}(\mathbf{G}) = \sum_i \sum_k \omega_k {\rho_i}^{ptc} g_{ik} \tag{37}$$

where ${\rho_i}^{ptc}$ is the tax credit to generator i from government in dollars per megawatt-hour. Production tax credits may be similarly applied to diurnal storage units.

Note that policies which provide a tax credit per unit of CO_2 sequestered through CCUS can be modeled either as a PTC with an appropriate tax credit rate, or as an emissions tax on e_i^{ccus} with a negative tax rate.

5.1.3 Investment Tax Credit

An investment tax credit transfers money from the government to electricity producers for newly built capacity of a specific type. These tax credits or subsidies effectively reduce the percieved investment cost of newly added capacity. Like the production tax credits, they therefore constitute a positive addition to the objective function. Investment tax credits are often established as a percentage of the total investment cost of new capacity, instead of a fixed amount. So, the total addition to the objective function (subtraction from the generator cost function) is:

$$C^{itc}_{POL}(\mathbf{A}) = \sum_i \rho^{itc}_i c^I_i a_i \tag{38}$$

where ρ^{itc}_i is the value of an investment tax credit as a percentage of investment costs. Investment tax credits may be similarly applied to diurnal storage units.

5.2 Policies Modeled as Constraints

Policy constraints are optional constraints that can be added to represent existing or proposed policies on any spatial level. Like with policy-related costs in the objective function, any number of these polices can be added.

5.2.1 Emission Cap

Emission caps are upper limits on the amount of emissions that can be emitted by generators in a certain area and time period. They can be regional or seasonal. Using the index π to refer to an emissions cap on pollutant $\mathfrak{p}_\pi$, the cap is represented as:

$$\sum_i \sum_k h_k I^\pi_{ik} e^{\mathfrak{p}}_i g_{ik} \leq E_\pi \quad \forall \pi \in \Pi \tag{39}$$

where I^π_{ik} is a binary indicator parameter to indicate whether emissions from generator i and hour k are included in the current constraint; E_π is the cap level, the maximum allowable emissions for cap policy π; and Π is the set of all emissions caps. Note that we use h_k instead of ω_k so that we can constrain the total amount of emissions without having to scale E_π to be an average hourly value.

5.2.2 RPS/CES Constraint

Many states in the US have implemented policies which mandate low- and non-emitting generation by way of a renewable portfolio standard (RPS), or its generalization, a clean electricity standard (CES). Both policy types at their core set a minimum on the amount of qualifying generation in an area. Typically, policies called RPSs require generation solely from renewable sources such as wind, solar, or geothermal. Policies called CESs tend to be more flexible, and can have a crediting system which assigns clean electricity credits (CECs) to generation based on its emission rate or other attributes. An RPS can be thought of as a CES in which all qualifying generation is renewable, and all qualifying generation equally earns one CEC per MWh of generation. The minimum amount of CECs that must be generated in an RPS/CES policy is typically defined as a percentage of the total electricity demand in an area, so the constraint is formulated as:

$$\sum_i \sum_k h_k \varphi^\phi_i g_{ik} \geq \sum_k h_k \Big(\sum_j \Omega^\phi_j \big(d_{jk} + \sum_{i \in \mathcal{G}^{\mathrm{DAC}}} I^j_i g_{ik} + \sum_b (1 - \xi_b) s_{bk} \big) \Big) \quad \forall \phi \in \Phi \tag{40}$$

where φ^ϕ_i is the rate at which generator i earns CECs toward policy ϕ, in CECs per MWh generated, and Ω^ϕ_j is the percentage of electricity demand at node j which must be covered with CECs.

DAC devices may also generate CECs to contribute to the RPS/CES requirement, and can be included with the appropriate crediting coefficient on the left hand side of constraint (40). However, notice that both the electricity used by DAC and the electricity lost during storage operation count towards the electricity demand in the policy area, and so are incorporated into the right hand side of constraint (40). The electricity loss from diurnal storage is written in the form derived in equation (23).

Even though they are not RPSs or CESs by name, utility decarbonization goals and estimated voluntary green power purchasing may be modeled as such.

5.2.3 Capacity Targets

Some policies set targets on the minimum (or rarely, maximum) amount of a certain type of generation that must be active. The sum of the active capacities for all the generators that belong to group m should be between the lower bound and upper bound of the capacity limits of group m:

$$A_m^- \leq \sum_i I_i^m (p_i^0 - r_i + a_i) \leq A_m^+ \quad \forall m \in \mathcal{M}_{pol} \tag{41}$$

where A_m^- is the minimum total capacity for capacity constraint m, A_m^+ is the maximum total capacity for capacity constraint m, and I_i^m is an indicator parameter which equals 1 if generator i is in capacity constraint m. This type of constraint can also be applied to diurnal storage capacity. For example, New York policy requires there to be 3000 MW of battery storage capacity in the state by 2030.

6 Dual Variables

Each constraint in any linear program, including E4ST, has a dual variable. The value of a constraint's dual variable in an optimal solution of a linear program is a measure of the change in the objective function value that would result from an infinitesimally small relaxation of the constraint, divided by the size of the relaxation. So, the units of the dual variable are the units of the objective function, divided by the units of the constraint. Several of E4ST's dual variables, which we will refer to as shadow prices, can be used to calculate interesting results about the optimal solution of the power system in E4ST.

6.1 Locational Marginal Prices

Taking a closer look at the nodal power balance constraints (30), we can see that the constraint is in units of MWhs, and that adding 1 to the right hand side of the constraint would result in having 1 more available MWh of electricity at that node, either by consuming one less or delivering one more. The shadow price on this constraint for node j and hour k, λ_{jk}, will then tell us the marginal value to the system-wide objective function of that MWh of electricity at node j during hour k. However, remember that the benefits and costs in each representative hour do not contribute equally to the objective function. An extra dollar of generator costs in hour k reduces the objective function by ω_k dollars, because that is the weight of the hour in the objective function. So, to convert the shadow price, which is a measure of effect on the objective function, back into an actual price for hour k we must divide the shadow price by the weights

$$\text{LMP}_{jk} = \frac{\lambda_{jk}}{\omega_k} \tag{42}$$

LMP_{jk} is then the locational marginal price of electricity at node j in hour k, in units of \$/MWh.

6.2 Policy Trading Credit Prices

We can also use the shadow prices on policy constraints for similarly useful results. For an emissions cap, we can use a carbon emissions cap as an example. The constraint (39) would be in units of tons of CO_2. However, unlike the nodal balance constraint, there is a single emissions cap for the entire year, so the shadow price on emissions cap π, λ_π, captures the average marginal abatement cost of CO_2 emissions. This can be used as a measure of the price of a CO_2 trading credit, in units of \$/ton of CO_2.

We can do something very similar for the other constraint-based policies, as these all have a single constraint for the entire year. The RPS/CES constraints are in units of clean electricity credits, so the shadow price for RPS/CES policy ϕ, λ_ϕ, is the market price of a policy credit.

References

[1] Biao Mao. "E4ST: Modeling, Simulation and Benefit-Cost Analysis Tools for Power System Planning and Policy Analysis". PhD thesis. Rensselaer Polytechnic Institute, Apr. 2017.

[2] R.D. Zimmerman and C. E. Murillo-Sánchez. *MATPOWER User's Manual, Version 7.1*. Aug. 2020. DOI: `10.5281/zenodo.4074122`. URL: `https://matpower.org/docs/MATPOWER-manual-7.1.pdf`.

[3] R.D. Zimmerman, C.E. Murillo-Sánchez, and R.J. Thomas. "MATPOWER: Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education". In: *IEEE Transactions on Power Systems* 26.1 (Feb. 2011), pp. 12–19. DOI: `10.1109/TPWRS.2010.2051168`.