Appendix to

"A computationally efficient Benders decomposition for energy systems planning problems with detailed operations and time-coupling constraints"

Anna Jacobson[1], Filippo Pecci [2], Nestor Sepulveda [3,4],Qingyu Xu [5] Jesse Jenkins [2,6]

1. Lewis-Sigler Institute of Integrative Genomics, Princeton University, Princeton, NJ, United States of America. email: annafj@princeton.edu.

2. Andlinger Center for Energy and the Environment, Princeton University, Princeton, NJ, United States of America

3. Research Affiliate, Massachusetts Institute of Technology, Cambridge, MA, Untied States of America

4. Manager, McKinsey and Company, United States of America

5. Energy Internet Research Institute, Tsinghua University, Beijing, People's Republic of China

6. Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, United States of America

A1 Notation

Table A1: A list of all sets included in the simulation.

Set	Composes	Set	Composes
G	All resources	G^{STOR}	All storage resources
G^{UC}	All resources subject to unit commitment	G^{HYDRO}	All hydro power resources
G^{NONRET}	All resources that cannot be retired	S	All consumer segments (demand)
G^{RPS}	All resources qualifying for RPS policy	L	All transmission lines
Z	All spatial zones	W	All subperiods
H_w	All hours per subperiod $w \in W$	I_z^{out}	Lines carrying power out of zone z
I_z^{in}	Lines carrying power into zone z		

Table A2: A list of all parameters in the simulation.

Param	Definition	Param	Definition
	Investments and Capacity		
$y_g^{P,0}$	Existing capacity [MW], resource g	$\overline{y}_g^P$	Max capacity [MW], resource g
$y_g^{E,0}$	Existing storage capacity [MWh], resource g	$\overline{y}_g^E$	Max storage capacity [MWh], resource g
$\gamma_g^{P,\text{SIZE}}$	Capacity size [MW]†, resource g	$\gamma_g^{E,\text{SIZE}}$	Storage capacity size [MWh]†, resource g
$y_l^{F,0}$	Existing transmission capacity [MW], line l	$\overline{y}_l^F$	Max transmission capacity [MW], line l
$\underline{\rho}_g$	Min duration, resource g^* [MWh/MW]	$\overline{\rho}_g$	Max duration, resource g^* [MWh/MW]
	Operations		
$\sigma_{g,t}$	Availability [%], resource g, timestep t	$d_{z,t}$	Net demand [MWh], zone z, timestep t
$\overline{\nu}_s$	Max non-served energy (NSE) [%], segment s	$\underline{\theta}_g$	Min output [%], resource g
β_w	Weight assigned to subperiod w [-]	δ_w	Number of hours in subperiod w [-]
α_t	Weight assigned to hour t [-]	ρ_g	Duration* for hydro resource g [MWh/MW]
η_g^+	Charging efficiency [%], storage resource g	η_g^-	Discharging efficiency [%], storage resource g
η_g^{disch}	Self-discharge rate from storage resource g [%]	$\sigma_{g,t}^{\text{inflow}}$	Norm. inflow, hydro resource g, timestep t [%]
μ_g^{up}	Max ramp up rate [%/hr], resource g	μ_g^{dn}	Max ramp down rate [%/hr], resource g
τ_g^{Up}	Min up time [hours], resource g	τ_g^{Dn}	Min down time [hours], resource g
	Policies		
$\underline{\epsilon}^{\text{RPS}}$	Share of demand in RPS constraint [%]	$\overline{\epsilon}^{\text{CO2}}$	CO2 emission cap [tons/MWh]
ϵ_g^{CO2}	CO2 emission factor [tons/MWh], resource g		
	Costs		
$c_g^{\text{P,INV}}$	Cost of investment in resource g [$/MW-yr]	$c_g^{\text{P,FOM}}$	Fixed O&M cost of resource g [$/MW-yr]
$c_g^{\text{E,INV}}$	Cost of investment, storage resource g [$/MWh-yr]	$c_g^{\text{E,FOM}}$	Fixed O&M cost, storage resource g [$/MWh-yr]
$c_l^{\text{F,INV}}$	Cost of investment in line l [$/MW-yr]	$c_{g,t}^{\text{VAR}}$	Variable costs [$/MWh], resource g, timestep t
$c_{s,z}^{\text{NSE}}$	Cost of NSE [$/MWh], segment s, zone z	c_g^{START}	Cost to start up resource $g \in G^{\text{UC}}$ [$]
c^{RPS}	Cost of RPS constraint noncompliance [$/MWh]	c^{CO2}	Cost of CO_2 constraint noncompliance [$/tons]

$*$ "Duration" here refers to the ratio between power and energy of a given resource, in MWh/MW

$\dagger$ Set to 1 for $g \notin G^{\text{UC}}$

Table A3: A list of all variables in the simulation.

Var.	Definition	Var.	Definition
	Investments and Capacity		
y_g^P	Capacity [MW], resource g	$y_g^{P,\text{NEW}}$	Investments in generation resource g^* [-]
y_g^E	Capacity [MWh], storage resource g	$y_g^{E,\text{NEW}}$	Investments in storage resource g^* [-]
y_l^F	Capacity* [MW], transmission line l	$y_l^{F,\text{NEW}}$	Investments in transmission line l^* [MW]
$y_g^{P,\text{RET}}$	Retirements, generation resource g^* [-]	$y_g^{E,\text{RET}}$	Retirements, storage resource g^* [-]
	Operations		
$x_{g,t}^{\text{inj}}$	Generation[MWh], resource g, timestep t	$x_{g,t}^{\text{wdw}}$	Withdrawals [MWh], of $g \in G^{\text{STOR}}$, timestep t
$x_{s,z,t}^{\text{nse}}$	NSE [MWh], segment s, zone z, timestep t	$x_{g,t}^{\text{soc}}$	State of charge [MWh] for $g \in G^{\text{STOR}}$, timestep t
$x_{g,t}^{\text{lvl}}$	Reservoir level [MWh] for $g \in G^{\text{HYDRO}}$, timestep t	$x_{g,t}^{\text{spill}}$	Spillage [MWh] from $g \in G^{\text{HYDRO}}$ in timestep t
$x_{l,t}^{\text{flow}}$	Flow [MWh] across line l in timestep t	$x_{g,t}^{\text{commit}}$	Units [-] of $g \in G^{\text{UC}}$ commited in timestep t
$x_{g,t}^{\text{shut}}$	Units [-] of $g \in G^{\text{UC}}$ shut down in timestep t [-]	$x_{g,t}^{\text{start}}$	Units [-] of $g \in G^{\text{UC}}$ started up in timestep t
	Policies		
x_w^{RPS}	Noncompliance [MWh] with RPS policy	x_w^{CO2}	Noncompliance [tons / MWh] with CO2 cap policy

$*$ Variable is included as integer in the MILP problem formulations

A2 Problem Description

To complement the constraints shown in the main text as a monolithic model, we present below an in-depth description of the separate groups of constraints in our model formulation.

For descriptions of variables, sets, and parameters used in sections A2.1, A2.2, A2.3, and A2.4, refer to the nomenclature table in section A1.

A2.1 Investment constraints

The following constraints model resource capacity given investments and retirements:

$$\gamma_g^{P,\text{SIZE}} y_g^{P,\text{NEW}} \le \overline{y}_g^P, \quad g \in G \tag{A1a}$$
$$\gamma_g^{P,\text{SIZE}} y_g^{P,\text{RET}} \le y_g^{P,0}, \quad g \in G \tag{A1b}$$
$$y_g^{P,\text{RET}} = 0, \quad g \in G^{NONRET} \tag{A1c}$$
$$y_g^P = y_g^{P,0} + \gamma_g^{P,\text{SIZE}} (y_g^{P,\text{NEW}} - y_g^{P,\text{RET}}), \quad g \in G, \tag{A1d}$$

Additional variables and constraints are included to allow for the modeling of energy storage resources. The following constraints model investments in storage:

$$\gamma_g^{E,\text{SIZE}} y_g^{E,\text{NEW}} \le \overline{y}_g^E, \quad g \in G^{\text{STOR}} \tag{A2a}$$
$$\gamma_g^{E,\text{SIZE}} y_g^{E,\text{RET}} \le y_g^{E,0}, \quad g \in G^{\text{STOR}} \tag{A2b}$$
$$y_g^{E,\text{RET}} = 0, \quad g \in G^{NONRET} \cap G^{\text{STOR}} \tag{A2c}$$
$$y_g^E = y_g^{E,0} + \gamma_g^{E,\text{SIZE}} (y_g^{E,\text{NEW}} - y_g^{E,\text{RET}}), \quad g \in G^{\text{STOR}} \tag{A2d}$$
$$\underline{\rho}_g y_g^P \le y_g^E, \quad g \in G^{\text{STOR}} \tag{A2e}$$
$$\overline{\rho}_g y_g^P \ge y_g^E, \quad g \in G^{\text{STOR}}, \tag{A2f}$$

Transmission capacity expansion is modeled by:

$$y_l^{F,\text{NEW}} \le \overline{y}_l^F, \quad l \in L \tag{A3a}$$
$$y_l^F = y_l^{F,0} + y_l^{F,\text{NEW}}, \quad l \in L, \tag{A3b}$$

A2.2 Operational constraints

System operation is primarily subject to the power demand balance stating that zonal demand equals available power within a zone (the sum of generation and net storage discharge and transmission imports) plus curtailed demand:

$$\sum_{g \in G_z} x_{g,t}^{\text{inj}} - \sum_{g \in G_z^{\text{STOR}}} x_{g,t}^{\text{wdw}} - \sum_{l \in I_z^{\text{out}}} x_{l,t}^{\text{flow}} + \sum_{l \in I_z^{\text{in}}} x_{l,t}^{\text{flow}} + \sum_{s \in S} x_{s,z,t}^{\text{nse}} = d_{z,t}, \quad z \in Z,\ t \in H_w,\ w \in W, \tag{A4}$$

Some resources have a constant available dispatchable capacity equal to their installed capacity, while other resources (e.g., VRE) may vary in maximum output throughout the planning period (e.g. based on wind or solar resource variability). To capture this, we introduce a maximum power output parameter $\sigma_{g,t} \in [0, 1]$ to model the fraction of installed power capacity available for dispatch at time $t \in H_w$. Note that for many resources (e.g., thermal and storage,) $\sigma_{g,t} = 1$ to denote the resource being consistently available at its full capacity. We also consider parameters ρ_g representing the ratio of power to energy for storage in hydropower resource $g \in G^{\text{HYDRO}}$, η_g^+ representing the charging efficiency for storage resource $g \in G^{\text{STOR}}$, and η_g^- representing the discharging efficiency for storage resource $g \in G^{\text{STOR}}$. Dispatched power and stored energy are limited by installed capacity, resulting in the following linear constraints:

$$x^{\text{inj}}_{g,t} \le \sigma_{g,t} y^P_g, \quad g \in G \setminus G^{\text{UC}},\ t \in H_w,\ w \in W \tag{A5a}$$
$$x^{\text{wdw}}_{g,t} \le \sigma_{g,t} y^P_g, \quad g \in G^{\text{STOR}},\ t \in H_w,\ w \in W \tag{A5b}$$
$$x^{\text{inj}}_{g,t} + x^{\text{wdw}}_{g,t} \le y^P_g, \quad g \in G^{\text{STOR}},\ t \in H_w,\ w \in W \tag{A5c}$$
$$x^{\text{soc}}_{g,t} \le y^E_g, \quad g \in G^{\text{STOR}},\ t \in H_w,\ w \in W \tag{A5d}$$
$$x^{\text{lvl}}_{g,t} \le \rho_g y^P_g, \quad g \in G^{\text{HYDRO}},\ t \in H_w,\ w \in W \tag{A5e}$$
$$\eta^+_g x^{\text{wdw}}_{g,t} \le y^E_g - x^{\text{soc}}_{g,t}, \quad g \in G^{\text{STOR}},\ t \in H_w,\ w \in W \tag{A5f}$$
$$\frac{x^{\text{inj}}_{g,t}}{\eta^-_g} \le x^{\text{soc}}_{g,t}, \quad g \in G^{\text{STOR}},\ t \in H_w,\ w \in W. \tag{A5g}$$

Some resources have a minimum required output level, which is enforced by:

$$x^{\text{inj}}_{g,t} \ge \underline{\theta}_g y^P_g, \quad g \in G \setminus (G^{\text{UC}} \cup G^{\text{STOR}} \cup G^{\text{HYDRO}}),\ t \in H_w,\ w \in W \tag{A6a}$$
$$x^{\text{inj}}_{g,t} + x^{\text{spill}}_{g,t} \ge \underline{\theta}_g y^P_g, \quad g \in G^{\text{HYDRO}},\ t \in H_w,\ w \in W. \tag{A6b}$$

Power flow between zones is bounded by transmission capacity:

$$x^{\text{flow}}_{l,t} \le y^F_l, \quad l \in L,\ t \in H_w,\ w \in W \tag{A7a}$$
$$-x^{\text{flow}}_{l,t} \le y^F_l, \quad l \in L,\ t \in H_w,\ w \in W. \tag{A7b}$$

The maximum curtailed demand in each consumer segment is constrained to be at most a fraction $\overline{\nu}_s \in [0,1]$ of the zonal demand:

$$x^{\text{nse}}_{s,z,t} \le \overline{\nu}_s d_{z,t}, \quad s \in S, z \in Z,\ t \in H_w,\ w \in W. \tag{A8}$$

Next, we introduce notation to model storage, UC and ramping limits within subperiods. The first time step in a subperiod is denoted by $t^0_w = (w-1)\delta_w + 1$, for all $w \in W$, while the last time step in a subperiod is $t_w = \delta_w w$, for all $w \in W$. The time steps that are not at the start of a subperiod are included in subset $H^0_w = H_w \setminus \{t^0_w\}$, for all $w \in W$. Storage and hydropower operational constraints are as follows:

$$x^{\text{soc}}_{g,t} - x^{\text{soc}}_{g,t-1} = \eta^+_g x^{\text{wdw}}_{g,t} - \frac{x^{\text{inj}}_{g,t}}{\eta^-_g} - \eta^{\text{disch}}_g x^{\text{soc}}_{g,t-1}, \quad g \in G^{\text{STOR}},\ t \in H^0_w,\ w \in W \tag{A9a}$$
$$x^{\text{soc}}_{g,t^0_w} - x^{\text{soc}}_{g,t_w} = \eta^+_g x^{\text{wdw}}_{g,t^0_w} - \frac{x^{\text{inj}}_{g,t^0_w}}{\eta^-_g} - \eta^{\text{disch}}_g x^{\text{soc}}_{g,t_w}, \quad g \in G^{\text{STOR}},\ w \in W \tag{A9b}$$
$$x^{\text{lvl}}_{g,t} - x^{\text{lvl}}_{g,t-1} = \sigma^{\text{inflow}}_{g,t} y^P_g - x^{\text{inj}}_{g,t} - x^{\text{spill}}_{g,t}, \quad g \in G^{\text{HYDRO}},\ t \in H^0_w,\ w \in W \tag{A9c}$$
$$x^{\text{lvl}}_{g,t^0_w} - x^{\text{lvl}}_{g,t_w} = \sigma^{\text{inflow}}_{g,t^0_w} y^P_g - x^{\text{inj}}_{g,t^0_w} - x^{\text{spill}}_{g,t^0_w}, \quad g \in G^{\text{HYDRO}},\ w \in W. \tag{A9d}$$

Ramping limits for thermal units are enforced within each subperiod. The difference in discharge from a resource g in two consecutive steps is constrained as follows:

$$x^{\text{inj}}_{g,t} - x^{\text{inj}}_{g,t-1} \le \mu^{\text{up}}_g y^P_g, \quad g \in G \setminus G^{\text{UC}},\ t \in H^0_w,\ w \in W \tag{A10a}$$
$$x^{\text{inj}}_{g,t-1} - x^{\text{inj}}_{g,t} \le \mu^{\text{dn}}_g y^P_g, \quad g \in G \setminus G^{\text{UC}},\ t \in H^0_w,\ w \in W \tag{A10b}$$
$$x^{\text{inj}}_{g,t^0_w} - x^{\text{inj}}_{g,t_w} \le \mu^{\text{up}}_g y^P_g, \quad g \in G \setminus G^{\text{UC}},\ w \in W \tag{A10c}$$
$$x^{\text{inj}}_{g,t_w} - x^{\text{inj}}_{g,t^0_w} \le \mu^{\text{dn}}_g y^P_g, \quad g \in G \setminus G^{\text{UC}},\ w \in W. \tag{A10d}$$

Resources in G^{UC} can commit an integer number of units at each time step. To avoid having integer operational variables, we relax integrality constraints on UC variables. This is a common strategy in the solution of energy systems planning problems (Lara et al. 2018, Li et al. 2022), as integer variables are expected to have tight linear relaxations for UC operations - see the optimality gaps reported in Table 2 in Li et al. (2022). Note that we do not relax integrality constraints on investment decision variables, whose integer representation is critical to fully capture economies of scale in generation and transmission expansion. UC is modeled as follows:

$$\gamma_g^{P,\text{SIZE}} x_{g,t}^{\text{commit}} \le y_g^P, \quad g \in G^{\text{UC}},\ t \in H_w,\ w \in W \tag{A11a}$$
$$\gamma_g^{P,\text{SIZE}} x_{g,t}^{\text{start}} \le y_g^P, \quad g \in G^{\text{UC}},\ t \in H_w,\ w \in W \tag{A11b}$$
$$\gamma_g^{P,\text{SIZE}} x_{g,t}^{\text{shut}} \le y_g^P, \quad g \in G^{\text{UC}},\ t \in H_w,\ w \in W \tag{A11c}$$
$$x_{g,t}^{\text{inj}} \ge x_{g,t}^{\text{commit}} \underline{\theta}_g \gamma_g^{P,\text{SIZE}}, \quad g \in G^{\text{UC}},\ t \in H_w,\ w \in W \tag{A11d}$$
$$x_{g,t}^{\text{inj}} \le x_{g,t}^{\text{commit}} \sigma_{g,t} \gamma_g^{P,\text{SIZE}}, \quad g \in G^{\text{UC}},\ t \in H_w,\ w \in W \tag{A11e}$$
$$x_{g,t}^{\text{commit}} - x_{g,t-1}^{\text{commit}} = x_{g,t}^{\text{start}} - x_{g,t}^{\text{shut}}, \quad g \in G^{\text{UC}},\ t \in H_w^0,\ w \in W \tag{A11f}$$
$$x_{g,t_w^0}^{\text{commit}} - x_{g,t_w}^{\text{commit}} = x_{g,t_w^0}^{\text{start}} - x_{g,t_w^0}^{\text{shut}}, \quad g \in G^{\text{UC}},\ w \in W. \tag{A11g}$$

Ramping limits on resources in $g \in G^{\text{UC}}$ are given by:

$$\begin{aligned} x_{g,t}^{\text{inj}} - x_{g,t-1}^{\text{inj}} \le \gamma_g^{P,\text{SIZE}} \mu_g^{\text{up}} (x_{g,t}^{\text{commit}} - x_{g,t}^{\text{start}}) + \gamma_g^{P,\text{SIZE}} \min(\sigma_{g,t}, \max(\underline{\theta}_g, \mu_g^{\text{up}})) x_{g,t}^{\text{start}} \\ - \gamma_g^{P,\text{SIZE}} \underline{\theta}_g x_{g,t}^{\text{shut}}, \quad g \in G^{\text{UC}},\ t \in H_w^0,\ w \in W \end{aligned} \tag{A12a}$$
$$\begin{aligned} x_{g,t-1}^{\text{inj}} - x_{g,t}^{\text{inj}} \le \gamma_g^{P,\text{SIZE}} \mu_g^{\text{dn}} (x_{g,t}^{\text{commit}} - x_{g,t}^{\text{start}}) + \gamma_g^{P,\text{SIZE}} \min(\sigma_{g,t}, \max(\underline{\theta}_g, \mu_g^{\text{dn}})) x_{g,t}^{\text{shut}} \\ - \gamma_g^{P,\text{SIZE}} \underline{\theta}_g x_{g,t}^{\text{start}}, \quad g \in G^{\text{UC}},\ h \in H_w^0,\ w \in W \end{aligned} \tag{A12b}$$
$$\begin{aligned} x_{g,t_w^0}^{\text{inj}} - x_{g,t_w}^{\text{inj}} \le \gamma_g^{P,\text{SIZE}} \mu_g^{\text{up}} (x_{g,t_w^0}^{\text{commit}} - x_{g,t_w^0}^{\text{start}}) + \gamma_g^{P,\text{SIZE}} \min(\sigma_{g,t_w^0}, \max(\underline{\theta}_g, \mu_g^{\text{up}})) x_{g,t_w^0}^{\text{start}} \\ - \gamma_g^{P,\text{SIZE}} \underline{\theta}_g x_{g,t_w^0}^{\text{shut}}, \quad g \in G^{\text{UC}}, w \in W \end{aligned} \tag{A12c}$$
$$\begin{aligned} x_{g,t_w}^{\text{inj}} - x_{g,t_w^0}^{\text{inj}} \le \gamma_g^{P,\text{SIZE}} \mu_g^{\text{dn}} (x_{g,t_w^0}^{\text{commit}} - x_{g,t_w^0}^{\text{start}}) + \gamma_g^{P,\text{SIZE}} \min(\sigma_{g,t_w^0}, \max(\underline{\theta}_g, \mu_g^{\text{dn}})) x_{g,t_w^0}^{\text{shut}} \\ - \gamma_g^{P,\text{SIZE}} \underline{\theta}_g x_{g,t_w^0}^{\text{start}}, \quad g \in G^{\text{UC}}, w \in W. \end{aligned} \tag{A12d}$$

In addition, committed resources may be required to stay online (resp. offline) for a minimum period of time τ^{Up} (resp. τ^{Dn}) before being shut down (resp. restarted). These constraints are modeled within each subperiod via definition of index sets Ω^{up} and Ω^{dn} in (A14), which are then used to incorporate constraints (A13), stating:

$$x_{g,t}^{\text{commit}} \ge \sum_{k \in \Omega_{g,w}^{\text{up}}(t)} x_{g,k}^{\text{start}} \tag{A13a}$$
$$x_{g,t}^{\text{commit}} + \sum_{k \in \Omega_{g,w}^{\text{dn}}(t)} x_{g,k}^{\text{shut}} \le \frac{y_g^P}{\gamma_g^{P,\text{SIZE}}}, \tag{A13b}$$

with

$$\begin{aligned} \Omega_{g,w}^{\text{up}}(t) &= \{\phi_{w,\tau_g^{\text{Up}}}(t), \phi_{w,\tau_g^{\text{Up}}-1}(t), \ldots, \phi_{w,1}(t), g_{w,0}(t) = t\}, \quad t \in H_w \\ \Omega_{g,w}^{\text{dn}}(t) &= \{\phi_{w,\tau_g^{\text{Dn}}}(t), \phi_{w,\tau_g^{\text{Dn}}-1}(t), \ldots, \phi_{w,1}(t), g_{w,0}(t) = t\}, \quad t \in H_w \end{aligned} \tag{A14}$$

where $\phi_{w,n}(t)$ corresponds to the time index that is n steps before t in H_w, where H_w is considered as a circular array of length δ_w. For example, if $\delta_w = 10\ \forall w$, we have that $H_2 = \{11, \ldots, 20\}$, and $\phi_{2,1}(11) = 20$.

A2.3 Policy constraints

A renewable portfolio standard enforces that a given share, $\underline{\epsilon}^{\text{RPS}}$, of the total demand is produced by qualifying resources, $G^{\text{RPS}} \subset G$. In order to avoid infeasiblities, we enforce RPS by assigning a heavy penalty to the violation of the minimum energy share constraint. Let $x_w^{RPS} \ge 0$ be a slack variable used to evaluate the level of non-compliance with RPS in subperiod $w \in W$. For RPS cases, we include the following constraint:

$$\sum_{w \in W} \left(\Big(\sum_{t \in H_w} \sum_{g \in G^{\text{RPS}}} \alpha_t x_{g,t}^{\text{inj}} \Big) + x_w^{RPS} \right) \ge \underline{\epsilon}^{\text{RPS}} \Big(\sum_{w \in W} \sum_{t \in H_w} \sum_{z \in Z} \alpha_t d_{z,t} \Big), \tag{A15a}$$

where α_t is the weight assigned to the time step t such that $\alpha_t = \frac{\beta_w}{\delta_w}$, for all $t \in H_w$ and $w \in W$, β_w is the weight associated to representative subperiod w by the timeseries clustering algorithm (i.e., the number of hours represented by subperiod w). Observe that $\alpha_t = 1 \; \forall \; t \in T$ for the full year when $|W| = 52$.

As a policy constraint in the CO2 scenario, consider a CO2 emissions cap. Analogously to the RPS constraint, we introduce slack variables $x_w^{CO2} \geq 0$ to evaluate the level of non-compliance with the CO2 emissions cap in subperiod $w \in W$. In the formulation of scenario CO2, we include the following constraint:

$$\sum_{w \in W} \left(\sum_{t \in H_w} \Big(\sum_{g \in G} \alpha_t \epsilon_g^{\text{CO2}} x_{g,t}^{\text{inj}} + \sum_{g \in G^{\text{STOR}}} \alpha_t \epsilon_g^{\text{CO2}} x_{g,t}^{\text{wdw}} \Big) - x_w^{CO2} \right) \leq \overline{\epsilon}^{\text{CO2}}, \tag{A16a}$$

where ϵ_g^{CO2} represents tons of CO2 emissions per-MWh for resources $g \in G$, and $\overline{\epsilon}^{\text{CO2}}$ is the maximum emission threshold. In our case, we define:

$$\overline{\epsilon}^{\text{CO2}} = 0.05 \Big(\sum_{w \in W} \sum_{t \in H_w} \sum_{z \in Z} \alpha_t d_{z,t} \Big). \tag{A17}$$

For scenario REF, CO2 emissions and generation portfolio are unrestricted. We effectively see $x_w^{RPS} = 0,\; x_w^{CO2} = 0,\; \underline{\epsilon}^{\text{RPS}} = 0,\; \overline{\epsilon}^{\text{CO2}} = \infty \quad \forall w \in W$.

A2.4 Objective function

Planning problems in energy systems minimize over costs of both investment and operations. For our purposes, we annualize investment costs and minimize one net cost function for the operations and investments over the year-long planning period.

Total fixed cost is given by:

$$\begin{aligned} c^{\text{FIXED}} &= \sum_{g \in G} c_g^{\text{P,INV}} \gamma_g^{P,\text{SIZE}} y_g^{P,\text{NEW}} + \sum_{g \in G^{\text{STOR}}} c_g^{\text{E,INV}} \gamma_g^{E,\text{SIZE}} y_g^{E,\text{NEW}} + \sum_{g \in G^{\text{HYDRO}}} c_g^{\text{E,INV}} \rho_g y_g^{P,\text{NEW}} \\ &+ \sum_{g \in G} c_g^{\text{P,FOM}} y_g^P + \sum_{g \in G^{\text{STOR}}} c_g^{\text{E,FOM}} y_g^E + \sum_{g \in G^{\text{HYDRO}}} c_g^{\text{E,FOM}} \rho_g y_g^P + \sum_{l \in L} c_l^{\text{F,INV}} y_l^{F,\text{NEW}}. \end{aligned} \tag{A18}$$

Next, denote by c_g^{VAR} resources' variable cost per MWh and compute total variable costs as:

$$c^{\text{VAR}} = \sum_{w \in W} \sum_{t \in H_w} \sum_{g \in G} c_g^{\text{VAR}} \alpha_t x_{g,t}^{\text{inj}} + \sum_{w \in W} \sum_{t \in H_w} \sum_{g \in G^{\text{STOR}}} c_g^{\text{VAR}} \alpha_t x_{g,t}^{\text{wdw}} \tag{A19}$$

Let $c_{s,z,t}^{\text{NSE}}$ be the cost of curtailing demand in consumer segment $s \in S$, zone $z \in Z$, and time $t \in H_w$ in subperiod $w \in W$. The total cost for curtailed demand, is:

$$c^{\text{NSE}} = \sum_{w \in W} \sum_{t \in H_w} \sum_{z \in Z} \sum_{s \in S} c_{s,z}^{\text{NSE}} \alpha_t x_{s,z,t}^{\text{nse}} \tag{A20}$$

Let c_g^{START} be cost to start-up a unit of resource $g \in G^{\text{UC}}$. The total start-up costs are given by:

$$c^{\text{START}} = \sum_{w \in W} \sum_{t \in H_w} \sum_{g \in G^{\text{UC}}} c_g^{\text{START}} \alpha_t x_{g,t}^{\text{start}} \tag{A21}$$

Finally, we include penalty costs for violating policy constraints:

$$c^{\text{RPS}} + c^{\text{CO2}} = \sum_{w \in W} c^{\text{RPS}} x_w^{RPS} + \sum_{w \in W} c^{\text{CO2}} x_w^{CO2}, \tag{A22}$$

where c^{RPS} is equal to the cost for violating the RPS policy, and c^{CO2} is the cost for violating the CO2 emissions cap. Total cost of the problem will be equal to:

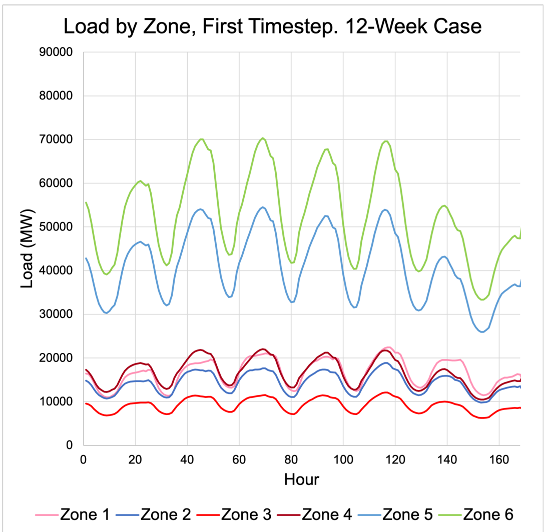
$$c^{\text{FIXED}} + c^{\text{VAR}} + c^{\text{NSE}} + c^{\text{START}} + c^{\text{RPS}} + c^{\text{CO2}}. \tag{A23}$$

A2.5 Input Data

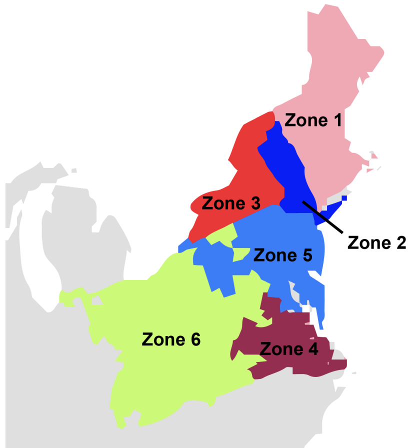
Data for PowerGenome comes from the Public Utility Data Liberation (PUDL) project and the Annual Technology Baseline (ATB) from the National Renewable Energy Lab (NREL) (Schivley et al. 2021). Demand and initial capacity for resources is identical across all CO2, RPS, and reference cases. Initial capacity and demand is also identical within each zone for varying spatial extent.

Figure A1: Load for the 6-zone, 12-week case. Shows load across all timesteps (A1c), load for the first week of the simulation (A1a), and a map of the zones included (A1b.)

(a) Load for the 6-zone 12-week case, first timestep



(b) Map of zones in the 6-zone case



(c) Load for the 6-zone 12-week case, all timesteps

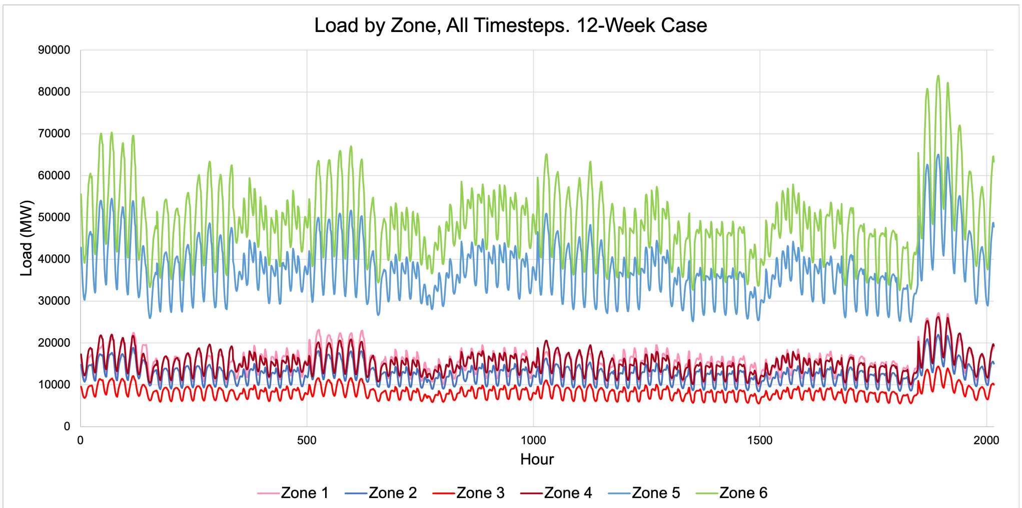
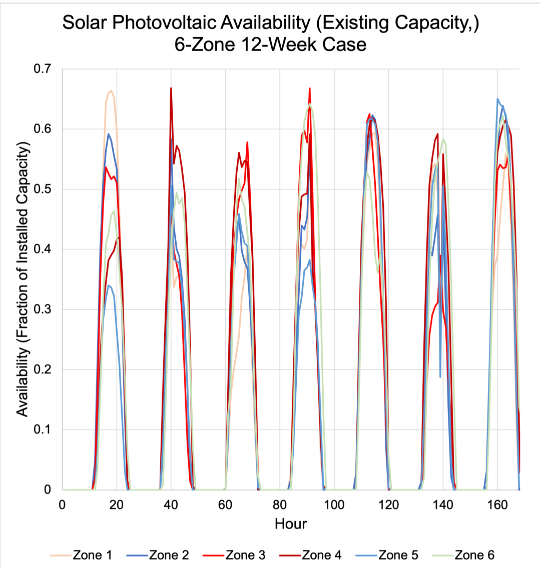
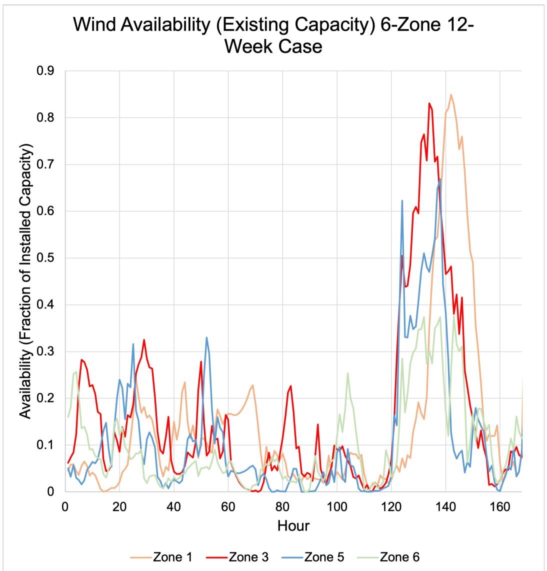


Figure A2: Availability of VRE for the 6-zone 12-week case. Figure shows the first timestep of the simulation. Zones correspond to the map in figure A1b.

(a) Availability of solar for existing photovoltaic capacity, by zone (see A1b). First timestep, 6-zone 12-week case.



(b) Availability of wind for existing wind capacity, by zone (see A1b). First timestep, 6-zone 12-week case.



A3 Results

We have included the LP relaxation gap in table A4. As noted on page 16 of the main text, trends in objective do not capture the the full implications of abstractions on models whose primary purpose is in output of variables' values and not simple system cost. As such, LP relaxation gap only tells a partial story of the impact of the transformation from LP to MILP formulations.

Donohoo-Vallett (2014) notes that in linearizing investment decisions in transmission, costs for expansion can only be considered on a per-MW basis. Models will always select the from the pool of the cheapest class of transmission line first, potentially inducing a bias towards smaller lines (Donohoo-Vallett 2014). This may prevent full capturing of economies of scale as an emergent property of systems. Further exploration of the specific per-generator impact of linearization on investment decisions is warranted but remains outside the scope of this manuscript.

$\|Z\|$	LP Relaxation Gap (%)		
	REF	RPS	CO2
2	6.81e-2	7.81e-2	7.73e-2
6	8.81e-2	7.97e-2	6.65e-2
12	9.61e-2	0.118	0.106
19	9.34e-2	0.114	0.115

Table A4: LP relaxation gap (%), formulated as gap in objective value between the LP and MILP case for each of the trials run. $100 \cdot (\mathrm{UB}_{\mathrm{MILP}} - \mathrm{LB}_{\mathrm{LP}})/\mathrm{LB}_{\mathrm{LP}}$.

Table A5: Runtime for Benders and monolithic models (100s,) followed by ratio between monolithic and decomposition solution approaches. ∞ denotes a case with an intractable monolithic model. Cases that are intractable due to memory are noted with the superscript M. Cases that are intractable due to insufficient time are noted with a superscript T. Cases where the model outperformed its analogous model formulation are shown for the runtime rows, cases where the decomposed model outperformed monolithic are bolded for the ratio cases. Results shown for the reference case. Additional cases are included in the appendix (Table A6) and in the main text (Table 3, section 4.2.)

	$\|Z\|$	$\|G\|$	LP						MILP					
	Weeks →		2	12	22	32	42	52	2	12	22	32	42	52
Benders (100s)	2	62	0.9	1.0	**1.2**	**1.3**	**1.5**	**1.8**	0.9	**1.0**	**1.2**	**1.3**	**1.5**	**1.6**
	6	175	3.4	6.9	10.7	**15.7**	**20.7**	**22.9**	4.1	**8.0**	**11.8**	**17.2**	**22.4**	**22.3**
	12	285	31.3	68.2	105.4	134.9	175.9	192.0	34.6	**89.0**	**115.9**	**158.0**	**210.0**	**196.4**
	19	437	231.0	408.9	522.0	748.4	772.1	900.7	194.1	**439.9**	**866.1**	**1107.5**	**1336.8**	**1235.9**
Mono. (100s)	2	62	**0.4**	**0.9**	1.7	3.0	3.7	5.8	**0.6**	5.6	6.5	13.7	28.4	41.1
	6	175	**0.6**	**4.4**	**9.5**	18.4	26.8	43.1	**1.3**	52.1	165.2	188.1	437.0	836.7
	12	285	**1.0**	**9.9**	**29.0**	47.7	75.3	102.8	**3.9**	154.4	950.0	1591.5	1888.6	∞^T
	19	437	**1.5**	**20.4**	**54.8**	**104.7**	**160.1**	**222.2**	**11.2**	660.0	∞^T	∞^T	∞^T	∞^T
Ratio	2	62	0.4	0.9	**1.4**	**2.3**	**2.5**	**3.2**	0.6	**5.3**	**5.6**	**10.3**	**18.7**	**25.5**
	6	175	0.2	0.6	0.9	**1.2**	**1.3**	**1.9**	0.3	**6.6**	**14.0**	**10.9**	**19.4**	**37.4**
	12	285	<0.1	0.1	0.3	0.3	0.4	0.6	0.1	**1.7**	**8.2**	**10.1**	**9.0**	$\boldsymbol{\infty^T}$
	19	437	<0.1	0.1	0.1	0.1	0.2	0.3	0.1	**1.5**	$\boldsymbol{\infty^T}$	$\boldsymbol{\infty^T}$	$\boldsymbol{\infty^T}$	$\boldsymbol{\infty^T}$

Table A6: Runtime for Benders and monolithic models (100s,) followed by ratio between monolithic and decomposition solution approaches. ∞ denotes a case with an intractable monolithic model. Cases that are intractable due to memory are noted with the superscript M. Cases that are intractable due to insufficient time are noted with a superscript T. Cases where the model outperformed its analogous model formulation are shown for the runtime rows, cases where the decomposed model outperformed monolithic are bolded for the ratio cases. Results shown for the RPS-constrained case. Additional cases are included in the appendix (Table A5) and in the main text (Table 3, section 4.2.)

	$\|Z\|$	$\|G\|$	LP						MILP					
	Weeks →		2	12	22	32	42	52	2	12	22	32	42	52
Benders (100s)	2	62	0.9	1.1	**1.4**	**1.6**	**1.9**	**2.3**	1.0	**1.3**	**1.5**	**1.9**	**2.2**	**2.2**
	6	175	3.7	8.3	13.8	**16.2**	**20.0**	**24.3**	3.6	**10.1**	**16.3**	**22.2**	**28.6**	**38.3**
	12	285	33.1	61.2	1.2e2	1.3e2	1.6e2	1.8e2	**36.2**	**82.9**	**1.4e2**	**1.8e2**	**2.2e2**	**2.2e2**
	19	437	1.6e2	3.8e2	5.1e2	6.3e2	8.7e2	9.4e2	1.7e2	**4.5e2**	**6.0e2**	**1.2e3**	**1.3e3**	**1.2e3**
Mono. (100s)	2	62	**0.4**	**1.0**	1.7	2.8	3.8	5.1	**0.6**	6.9	10.2	20.6	44.4	2.1e2
	6	175	**0.6**	**3.7**	**9.3**	17.4	25.2	32.1	**2.2**	28.1	∞^T	3.1e2	4.4e2	8.7e2
	12	285	**1.0**	**9.7**	**24.8**	**47.5**	**76.3**	**1.1e2**	37.9	∞^T	∞^T	∞^T	∞^T	∞^T
	19	437	**1.7**	**21.8**	**55.3**	**1.0e2**	**1.6e2**	**2.3e2**	**9.8**	∞^T	∞^T	∞^T	∞^T	∞^T
Ratio	2	62	0.5	0.9	**1.3**	**1.8**	**2.0**	**2.3**	0.6	**5.5**	**6.7**	**11.1**	**20.5**	**97.3**
	6	175	0.2	0.4	0.7	**1.1**	**1.3**	**1.3**	0.6	**2.8**	$\boldsymbol{\infty^T}$	**14.2**	**15.3**	**22.7**
	12	285	<0.1	0.2	0.2	0.4	0.5	0.6	1.0	$\boldsymbol{\infty^T}$	$\boldsymbol{\infty^T}$	$\boldsymbol{\infty^T}$	$\boldsymbol{\infty^T}$	$\boldsymbol{\infty^T}$
	19	437	0.1	0.1	0.1	0.2	0.2	0.2	<0.1	$\boldsymbol{\infty^T}$	$\boldsymbol{\infty^T}$	$\boldsymbol{\infty^T}$	$\boldsymbol{\infty^T}$	$\boldsymbol{\infty^T}$

References

Donohoo-Vallett PE (2014) *Design of wide-area electric transmission networks under uncertainty: Methods for dimensionality reduction*. Ph.D. thesis, Massachusetts Institute of Technology.

Lara CL, Mallapragada DS, Papageorgiou DJ, Venkatesh A, Grossmann IE (2018) Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm. *European Journal of Operational Research* 271(3):1037–1054.

Li C, Conejo AJ, Liu P, Omell BP, Siirola JD, Grossmann IE (2022) Mixed-integer linear programming models and algorithms for generation and transmission expansion planning of power systems. *European Journal of Operational Research* 297(3):1071–1082.

Schivley G, Welty E, Patankar N (2021) Powergenome/powergenome: v0. 4.1.