Supplementary information:

Learning Spatio-Temporal Aggregations for Large-Scale Capacity Expansion Problems

Aron Brenner[1], Rahman Khorramfar[2], Saurabh Amin[3]

[1],[3] Civil and Environmental Engineering (CEE) and Laboratory for Information & Decision Systems (LIDS)
[2] MIT Energy Initiative (MITEI) and Laboratory for Information & Decision Systems (LIDS)
{khorram, abrenner, amins, dharik}@mit.edu

1 Model Formulation

Our formulation is based on the formulation proposed in [1], but we applied a set of assumption to simplify the model and speed up the running time. The model determines minimum cost investment and operational decision for power and NG system. across a set of representative periods. The formulation allows different temporal resolutions for the operation of both systems as data availability or planning requirements can be different for power and NG systems. The operations of both systems are coupled through two sets of constraints. The first set ensure NG flow to the power system. The second coupling constraints limit the CO_2 emission incurred by consuming NG in both power and NG systems.

The network consists of three sets of nodes as depicted in Figure 1. The first set represents power system nodes and is characterized by different generation technologies (plant types), demand, storage, and the set of adjacent nodes. The second set of nodes are NG nodes each of which associated with injection amount, demand, and its adjacent nodes. Storage tanks, vaporization and liquefaction facilities, which are commonly used in the non-reservoir storage of NG, collectively form the third set of nodes namely SVL nodes. The model also consider the *renewable natural gas* (RNG) which is type of net-zero biofuel fully interchangeable with NG and hence can be imported and transported by the NG pipelines [2]. Details regarding input data including generation plant and

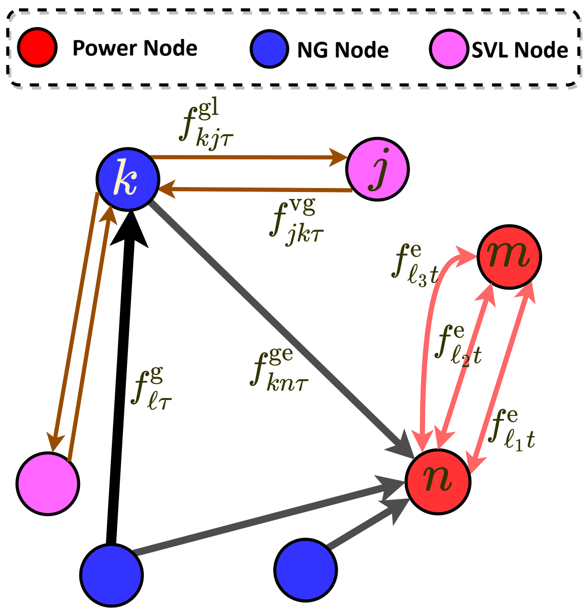

Figure 1: Power nodes can be connected by multiple bi-directional transmission lines denoted by $f^{\mathrm{e}}_{\ell t}$. Each power operate local NG-fired plants by drawing gas from nodes that are connected to it. The variable $f^{\mathrm{ge}}_{\ell\tau}$ captures this flow. Each NG node is connected to its adjacent SVL nodes through two unidirectional pipelines where one is from NG to SVL's liquefaction facilities denoted by $f^{\mathrm{gl}}_{kj\tau}$; and the other one from SVL's vaporization facility to NG node denoted by $f^{\mathrm{vg}}_{jk\tau}$. The variable $f^{\mathrm{g}}_{\ell\tau}$ denotes the flow between NG nodes. Similar to the power system, NG nodes can be connected by one or more uni-directional pipelines, but only one connection is depicted here. Candidate transmission lines and pipelines are not shown in this figure.

storage types, demand, and cost assumptions are provided in the Appendix of Supplementary Information of [1].

Indices	
n, m	Power system node
k	NG system node
j	SVL facility node
i	Power generation plant type
r	Storage type for power network
ℓ	Transmission line or pipeline
t	Time step for power system planning
τ	Time step for NG system planning

Sets	
$\mathcal{N}^{\mathrm{e}}$	Power system nodes
$\mathcal{P}$	Power plant types
$\mathcal{R} \subset \mathcal{P}$	VRE power plant types
$\mathcal{G} \subset \mathcal{P}$	gas-fired plant types
$\mathcal{H} \subset \mathcal{P}$	Thermal plant types
$\mathcal{T}^{\mathrm{e}}$	Representative hours for power system
$\mathfrak{R}$	Representative days
$\mathfrak{T}^{\mathrm{e}}_{\tau}$	Hours in the representative day τ
$\mathcal{L}^{\mathrm{e}}$	Existing and candidate transmission lines
$\mathcal{S}^{\mathrm{e}}_{n}$	Storage facility types
$\mathcal{A}^{\mathrm{g}}_{n}$	Adjacent NG nodes for node n
$\mathcal{L}^{\mathrm{e}}_{nm}$	Existing and candidate transmission lines between node n and m
$\mathcal{N}^{\mathrm{g}}, \mathcal{N}^{s}$	NG and SVL nodes
$\mathcal{T}^{\mathrm{g}}$	Days of the planning year for NG system
$\mathcal{A}^{s}_{k}$	Adjacent SVL facilities of node k
$\mathcal{L}^{\mathrm{g}}$	Existing and candidate pipelines
$\mathcal{L}^{\mathrm{gExp}}_{k}$	Existing and candidate pipelines starting from node k
$\mathcal{L}^{\mathrm{gImp}}_{k}$	Existing and candidate pipelines ending at node k

Annualized Cost Parameters	
C^{inv}_{i}	CAPEX of plants, [\$/plant]
C^{dec}_{i}	Plant decommissioning cost, [\$/plant]
$C^{\mathrm{trans}}_{\ell}$	Transmission line establishment cost, [\$/line]
C^{EnInv}_{r}	Storage establishment energy-related cost, [\$/MWh]
C^{pInv}_{r}	Storage establishment power-related cost, [\$/MW]
C^{pipe}_{ℓ}	Pipelines establishment cost, [\$/line]
C^{strInv}_{j}	CAPEX of storage tanks at SVLs, [\$/MMBtu]
C^{vprInv}_{j}	CAPEX of vapor. plants at SVLs, [\$/MMBtu]

Annual Costs

C_i^{fix}	FOM for plants, [\$]
C_r^{EnFix}	Energy-related FOM for storage, [\$/MWh]
C_r^{pFix}	Power-related FOM for storage, [\$/MW]
C_j^{strFix}	FOM for storage tanks, [\$/MMBtu]
C_j^{vprFix}	FOM for vaporization plants, [\$/MMBtu]

Other Cost Parameters

C_i^{var}	VOM for plants, [\$/MWh]
C^{eShed}	Unsatisfied power demand cost, [\$/MWh]
C_i^{fuel}	Fuel price for plants, [\$/MMBtu]
C^{ng}	Fuel price for NG, [\$/MMBtu]
C^{rng}	Price of RNG, [\$/MMBtu]
C^{gShed}	Unsatisfied NG demand cost [\$/MMBtu]

Other Parameters for the Power System

ρ_{nti}	Capacity factor
D_{nt}^{e}	Power demand, [MWh]
h_i	Heat rate, [MMBtu/MWh]
η_i	Carbon capture rate, [%]
U_i^{prod}	Nameplate capacity, [MW]
L_i^{prod}	Minimum stable output, [%]
U_i^{ramp}	Ramping limit, [%]
$\gamma_r^{\text{eCh}}, \gamma_r^{\text{eDis}}$	Charge/discharge rate for storage
I_ℓ^{trans}	Initial capacity for transmission line ℓ, [MW]
U_ℓ^{trans}	Upper bound for transmission line ℓ, [MW]
$\mathcal{I}_\ell^{\text{trans}}$	1, if trans. line ℓ exists; 0, otherwise
I_{ni}^{num}	Initial number of plants
$U_{\text{emis}}^{\text{e}}$	Baseline emission of CO_2 in 1990 from generation consumption, [ton]
L^{RPS}	Renewable Portfolio Standard (RPS) value
$\mathfrak{D}_n$	Distance between node n and CO_2 storage site
E^{pipe}	Electric requirement for CO_2 pipeline operations [MWh/mile/ton/hour]
E^{pump}	Electric requirement for compression of CO_2
E^{cprs}	Number of compressors required in the pipeline from node n to the storage site pipelines [MWh/ton/hour]
$U_{\mathcal{Q}}^{\text{prod}}$	Production capacity for set of plants $\mathcal{Q} \subset \mathcal{P}$, [MW]
ζ	Emission reduction goal
w_t	Number of periods represented by period t
ϕ_t^{e}	Mapping of representative period t to its original period in the time series

Other Parameters for the NG System	
$D^{\mathrm{g}}_{k\tau}$	NG demand, [MMBtu]
η^{g}	Emission factor for NG [ton CO_2/MMBtu]
U^{inj}_{k}	Upper bound for NG supply, [MMBtu]
$\gamma^{\mathrm{liqCh}}_{j}$	Charge efficiency of liquefaction plant
$\gamma^{\mathrm{vprDis}}_{j}$	Discharge efficiency of vaporization plant
β	Boil-off gas coefficient
I^{pipe}_{ℓ}	Initial capacity for pipeline ℓ, [MMBtu]
U^{pipe}_{ℓ}	Upper bound capacity for pipeline ℓ, [MMBtu]
$\mathcal{I}^{\mathrm{pipe}}_{\ell}$	1, if the pipeline ℓ exists; 0, otherwise
I^{gStr}_{j}	Initial storage capacity, [MMBtu]
I^{vpr}_{j}	Initial vaporization capacity, [MMBtu/d]
I^{liq}_{j}	Initial liquefaction capacity, [MMBtu/d]
I^{store}_{kj}	Initial capacity of storage facility
$U^{\mathrm{g}}_{\mathrm{emis}}$	Baseline emission of CO_2 in 1990 from non-generation consumption, [ton]
Ω_{τ}	Set of days represented by day $\tau \in \mathfrak{R}$

Investment Decision Variables	
$x^{\mathrm{op}}_{ni} \in \mathbb{Z}^{+}$	Number of available plants
$x^{\mathrm{est}}_{ni} \in \mathbb{Z}^{+}$	Number of new plants established
$x^{\mathrm{dec}}_{ni} \in \mathbb{Z}^{+}$	Number decommissioned plants
$y^{\mathrm{eCD}}_{nr} \in \mathbb{R}^{+}$	Charge/discharge capacity of storage battery
$y^{\mathrm{eLev}}_{nr} \in \mathbb{R}^{+}$	Battery storage level
$z^{\mathrm{e}}_{\ell} \in \mathbb{B}$	1, if transmission line ℓ is built; 0, otherwise
$z^{\mathrm{g}}_{\ell} \in \mathbb{B}$	1, if pipeline ℓ is built; 0, otherwise

1.1 Power System Model

Objective Function:

$$\min \sum_{n\in\mathcal{N}^{\mathrm{e}}} \sum_{i\in\mathcal{P}} (C^{\mathrm{inv}}_{i} x^{\mathrm{est}}_{ni} + C^{\mathrm{fix}}_{i} x^{\mathrm{op}}_{ni} + \sum_{r\in\mathcal{S}^{\mathrm{e}}_{n}} (C^{\mathrm{pInv}}_{r} + C^{\mathrm{pFix}}) y^{\mathrm{eCD}}_{nr}) +$$

$$\sum_{n\in\mathcal{N}^{\mathrm{e}}} \sum_{r\in\mathcal{S}^{\mathrm{e}}_{n}} (C^{\mathrm{EnInv}}_{r} + C^{\mathrm{EnFix}}) y^{\mathrm{eLev}}_{nr} + \tag{1a}$$

$$\sum_{n\mathcal{N}^{\mathrm{e}}} \sum_{i\in\mathcal{P}} C^{\mathrm{dec}}_{i} x^{\mathrm{dec}}_{ni} \tag{1b}$$

$$\sum_{n\in\mathcal{N}^{\mathrm{e}}} \sum_{i\in\mathcal{P}} \sum_{t\in\mathcal{T}^{\mathrm{e}}} w_t p_{nti} C^{\mathrm{var}}_{i} + \tag{1c}$$

$$\sum_{l\in\mathcal{L}^{\mathrm{e}}} C^{\mathrm{trans}}_{\ell} z^{\mathrm{e}}_{\ell} + \tag{1d}$$

$$\sum_{n\in\mathcal{N}^{\mathrm{e}}} \mathfrak{D}_n C^{\mathrm{inv}}_{\mathrm{CO_2}} \kappa^{\mathrm{pipe}}_{n} + C^{\mathrm{str}}_{\mathrm{CO_2}} \sum_{n\in\mathcal{N}^{\mathrm{e}}} \sum_{t\in\mathcal{T}^{\mathrm{e}}} w_t \kappa^{\mathrm{capt}}_{nt} + \tag{1e}$$

$$\sum_{n\in\mathcal{N}^{\mathrm{e}}} \sum_{i\in\mathcal{P}} \sum_{t\in\mathcal{T}^{\mathrm{e}}} w_t p_{nti} (C^{\mathrm{fuel}}_{i} h_i) + \tag{1f}$$

$$\sum_{n\in\mathcal{N}^{\mathrm{e}}} \sum_{t\in\mathcal{T}^{\mathrm{e}}} w_t C^{\mathrm{eShed}}_{n} a^{\mathrm{e}}_{nt} + \tag{1g}$$

Other Decision Variables for Power System	
$p_{nti} \in \mathbb{R}^+$	Generation rate, [MW]
$f^{\mathrm{e}}_{\ell t} \in \mathbb{R}$	Flow rates, [MW]
$s^{\mathrm{eCh}}_{ntr}, s^{\mathrm{eDis}}_{ntr} \in \mathbb{R}^+$	Storage charged/discharged, [MW]
$s^{\mathrm{eLev}}_{ntr} \in \mathbb{R}^+$	Storage level, [MWh]
$\kappa^{\mathrm{capt}}_{nt} \in \mathbb{R}^+$	Captured CO_2 [ton/h]
$\kappa^{\mathrm{pipe}}_{n} \in \mathbb{R}^+$	CO_2 pipeline capacity [ton/h]
$a^{\mathrm{e}}_{nt} \in \mathbb{R}^+$	Amount of load shedding, [MWh]
$\mathcal{E}^{\mathrm{e}}$	Total emission from power system

Other Decision Variables for NG System (all in MMBtu)	
$x^{\mathrm{gStr}}_{j} \in \mathbb{R}^+$	Installed additional storage capacities
$x^{\mathrm{vpr}}_{j} \in \mathbb{R}^+$	Installed additional vaporization capacities
$f^{\mathrm{g}}_{\ell\tau} \in \mathbb{R}^+$	Flow rates
$f^{\mathrm{ge}}_{kn\tau} \in \mathbb{R}^+$	Flow rates from NG nodes to power nodes
$f^{\mathrm{gl}}_{kj\tau} \in \mathbb{R}^+$	Flow rates from node NG nodes to liquefaction plants
$f^{\mathrm{vg}}_{jk\tau} \in \mathbb{R}^+$	Flow rates from vaporization plants to NG nodes
$g_{k\tau} \in \mathbb{R}^+$	NG supply (injection)
$s^{\mathrm{gStr}}_{j\tau} \in \mathbb{R}^+$	Storage capacities
$s^{\mathrm{vpr}}_{j\tau}, s^{\mathrm{liq}}_{j\tau} \in \mathbb{R}^+$	Vaporization and liquefaction amounts
$a^{\mathrm{g}}_{k\tau} \in \mathbb{R}^+$	Amount of load shedding
$a^{\mathrm{rng}}_{k\tau} \in \mathbb{R}^+$	Amount of RNG consumption
$\mathcal{E}^{\mathrm{g}}$	Total emission from NG system

The objective function (1) minimizes the total investment and operating costs incurred in power system. The first term (1a) is the investment and fixed operation and maintenance (FOM) costs for generation and storage. The term (1b) captures the cost of plant retirement or decommissioning. The variable operating and maintenance (VOM) are represented by term (1c). The network expansion costs is included by term (1d). The cost of CO2 transport and storage infrastructure required to accompany CC-CCS power generation is incorporated by term (1e) where it captures the cost associated with establishing CO_2 pipelines and storage. Note that we conservatively assume that each CO_2 pipeline connects a power node to the storage site, which ignores the possibility of meshed network design for CO2 transport. The cost of fuel consumption for non-gas-fired power plants (i.e., nuclear plant) are ensured by term (1f). The term term (1g) penalizes the load shedding in the power system which can occur due to unsatisfied demand.

Investment and Unit Commitment: For every $n \in \mathcal{N}^{\mathrm{e}}, i \in \mathcal{P}$

$$x^{\mathrm{op}}_{ni} = I^{\mathrm{num}}_{ni} - x^{\mathrm{dec}}_{ni} + x^{\mathrm{est}}_{ni} \tag{2a}$$

$$\tag{2b}$$

constraints (2a) specify the number of operating plants. The unit commitment constraints for each node and plant type are presented in (**??**) which computes the number of plants committed, started up, or shut down during a period. Constraints (**??**) limits the number of committed units to the number of available ones. As per other similar studies [**?** 4, 3**?**], we relaxed the integrality of unit commitment decisions to ease the computational complexity of the problem .

Generation, Ramping, and Load Shedding: For every $n \in \mathcal{N}^{\mathrm{e}}, t \in \mathcal{T}^{\mathrm{e}}$

$$L_i^{\mathrm{prod}} U_i^{\mathrm{prod}} x_{ni}^{\mathrm{op}} \leq p_{nti} \leq U_i^{\mathrm{prod}} x_{ni}^{\mathrm{op}} \qquad i \in \mathcal{H} \tag{3a}$$

$$|p_{nti} - p_{n,(t-1),i}| \leq U_i^{\mathrm{ramp}} U_{ni}^{\mathrm{prod}} x_{ni}^{\mathrm{op}} + \max(L_i^{\mathrm{prod}}, U_i^{\mathrm{ramp}}) U_i^{\mathrm{prod}} x_{ni}^{\mathrm{op}} \qquad i \in \mathcal{H} \tag{3b}$$

$$p_{nti} \leq \rho_{nti} U_{ni}^{\mathrm{prod}} x_{ni}^{\mathrm{op}} \qquad i \in \mathcal{R} \tag{3c}$$

$$a_{nt}^{\mathrm{e}} \leq D_{n\phi_t^{\mathrm{e}}}^{\mathrm{e}} \tag{3d}$$

the generation limits are imposed in constraints (3a). Constraints (3b) are the ramping constraints that limit the generation difference of thermal units in any consecutive time periods to a ramping limit in the right-hand-side of the equation. The generation pattern of VREs is determined by their hourly profile in the form of capacity factor; constraints (3c) limit the generation of VRE to hourly capacity factor (i.e. ρ_{nti}) of maximum available capacity (i.e. $U_{ni}^{\mathrm{prod}} x_{ni}^{\mathrm{op}}$). Constraints (3d) state that the load shedding amount can not exceed demand.

Power Balance Constraints: For every $n \in \mathcal{N}^{\mathrm{e}}, t \in \mathcal{T}^{\mathrm{e}}$

$$\sum_{i \in \mathcal{P}} p_{nti} + \sum_{m \in \mathcal{N}^{\mathrm{e}}} \sum_{l \in \mathcal{L}_{nm}^{\mathrm{e}}} sgn(n-m) f_{\ell t}^{\mathrm{e}} + \sum_{r \in \mathcal{S}_n^{\mathrm{e}}} (s_{ntr}^{\mathrm{eDis}} - s_{ntr}^{\mathrm{eCh}}) + a_{nt}^{\mathrm{e}} = D_{n\phi_t^{\mathrm{e}}}^{\mathrm{e}} + \mathfrak{D}_n E^{\mathrm{pipe}} \kappa_n^{\mathrm{pipe}} + E^{\mathrm{cprs}} E^{\mathrm{pump}} \kappa_{nt}^{\mathrm{capt}} \tag{4a}$$

constraints (4a) ensures that for each node at each planning period the generation, power curtailment amount, the net flow, and the net storage power is equal to the net demand. The net demand is defined in the right-hand-side where the first term is the baseline demand, the second term is the electricity consumption by CO_2 pipelines and the last term is the electricity used by compressors. The notation $sgn(n-m)$ is the *sign* function that takes value -1 if $n < m$, value 1 if $n > m$, and 0 otherwise. We use this function to ensure that $f_{\ell t}^{e}$ appears with opposite signs in the balance equations of the nodes connected by transmission line ℓ.

Network Constraints: For every $l \in \mathcal{L}^{\mathrm{e}}, t \in \mathcal{T}^{\mathrm{e}}, n, m \in \mathcal{N}_\ell^{\mathrm{e}}$

$$|f_{lt}^{\mathrm{e}}| \leq I_\ell^{\mathrm{trans}} \qquad \text{if } \mathcal{I}_\ell^{\mathrm{trans}} = 1 \tag{5a}$$

$$|f_{lt}^{\mathrm{e}}| \leq U_\ell^{\mathrm{trans}} z_\ell^{\mathrm{e}} \qquad \text{if } \mathcal{I}_\ell^{\mathrm{trans}} = 0 \tag{5b}$$

$$\tag{5c}$$

Flow for the existing transmission lines is limited by constraints (5a). Constraints (5b) limits the flow in candidate transmission lines only if it is already

established (i.e., z_ℓ^{t}=1). DC power flow constraints for candidate and existing transmission lines are respectively imposed in constraints (**??**) and (**??**). Constraints (**??**) restricts the phase angles, and constraints (**??**) sets the phase angle for the reference node 0.
Storage Constraints: For every $n \in \mathcal{N}^{\mathrm{e}}, t \in \mathcal{T}^{\mathrm{e}}, r \in \mathcal{S}_n^{\mathrm{e}}$

$$s_{ntr}^{\mathrm{eLev}} = s_{n,t-1,r}^{\mathrm{eLev}} + \gamma_r^{\mathrm{eCh}} s_{ntr}^{\mathrm{eCh}} - \frac{s_{ntr}^{\mathrm{eDis}}}{\gamma_r^{\mathrm{eDis}}} \tag{6a}$$

$$s_{ntr}^{\mathrm{eDis}} \leq y_{nr}^{\mathrm{eCD}}, s_{ntr}^{\mathrm{eCh}} \leq y_{nr}^{\mathrm{eCD}} \tag{6b}$$

$$s_{ntr}^{\mathrm{eLev}} \leq y_{nr}^{\mathrm{eLev}} \tag{6c}$$

$$s_{n,t_1,r}^{\mathrm{eLev}} = s_{n,t_{24},r}^{\mathrm{eLev}} \qquad t_1, t_{24} \in \mathfrak{T}_\tau^{\mathrm{e}}, \tau \in \mathfrak{R} \tag{6d}$$

Constraints (6a) model battery storage dynamics. The charge/discharge limits are imposed in (6b), and constraints (6c) limits the storage level. Note that as for other similar studies [4, 3], we do not account for storage capacity degradation. Representative days are not necessarily consecutive, therefore the formulation should account for the carryover storage level between representative days. Li et al. [3] enforce the beginning and ending storage levels of each representative days to 50% of the maximum storage level. In constraints (6d), we use a similar technique, yet more flexible, as we assume that beginning (i.e., t_1) and ending (i.e., t_{24}) storage levels are the same for any representative day.
Renewable Portfolio Standards (RPS):

$$\sum_{n \in \mathcal{N}^{\mathrm{e}}} \sum_{t \in \mathcal{T}^{\mathrm{e}}} \sum_{i \in \mathcal{R}} p_{nti} \geq L^{\mathrm{RPS}} \sum_{n \in \mathcal{N}^{\mathrm{e}}} \sum_{t \in \mathcal{T}^{\mathrm{e}}} D_{n\phi_t^{\mathrm{e}}}^{\mathrm{e}} \tag{7a}$$

The formulation requires the model to procure a certain share of the total demand from renewable energy sources. The share of renewable energy sources which is known as Renewable Portfolio Share (RPS) is imposed by constraint (7a).

1.2 NG System Model

Objective Function:

$$\min \sum_{l \in \mathcal{L}^{\mathrm{g}}} C_\ell^{\mathrm{pipe}} z_\ell^{\mathrm{g}} + \tag{8a}$$

$$\sum_{k \in \mathcal{N}^{\mathrm{g}}} \sum_{\tau \in \mathcal{T}^{\mathrm{g}}} C^{\mathrm{ng}} g_{k\tau} + \tag{8b}$$

$$\sum_{j \in \mathcal{N}^s} (C_j^{\mathrm{strInv}} x_j^{\mathrm{gStr}} + C_j^{\mathrm{vprInv}} x_j^{\mathrm{vpr}}) + \tag{8c}$$

$$\sum_{j \in \mathcal{N}^s} \Big(C^{\mathrm{strFix}} (I_j^{\mathrm{gStr}} + x_j^{\mathrm{gStr}}) + C^{\mathrm{vprFix}} (I_j^{\mathrm{vpr}} + x_j^{\mathrm{vpr}}) \Big) + \tag{8d}$$

$$\sum_{k \in \mathcal{N}^{\mathrm{g}}} \sum_{t \in \mathcal{T}^{\mathrm{g}}} (C^{\mathrm{rng}} a_{k\tau}^{\mathrm{rng}} + C^{\mathrm{gShed}} a_{k\tau}^{\mathrm{ng}}) \tag{8e}$$

The objective function (8) minimizes the total investment and operating costs incurred in the NG system. The first term (8a) is the investment cost for establishing new pipelines. The second term (8b) is the cost of procuring NG from various sources to the system. For example, New England procures its NG from Canada, and its adjacent states such as New York. The term (8c) and (8d) handle the investment and FOM costs associated with NG storage, respectively. The last term (8e) captures the cost of using RNG and NG load shedding.

NG Balance Constraint: For every $k \in \mathcal{N}^{\mathrm{g}}, \tau \in \mathcal{T}^{\mathrm{g}}$

$$
\begin{aligned}
& g_{k\tau} - \sum_{l \in \mathcal{L}_k^{\mathrm{gExp}}} f_{\ell\tau}^{\mathrm{g}} + \sum_{l \in \mathcal{L}_k^{\mathrm{gImp}}} f_{\ell\tau}^{\mathrm{g}} - \sum_{n \in \mathcal{A}_k^{\mathrm{e}}} f_{kn\tau}^{\mathrm{ge}} + \\
& \sum_{j \in \mathcal{A}_k^{s}} (f_{jk\tau}^{\mathrm{vg}} - f_{kj\tau}^{\mathrm{gl}}) + a_{k\tau}^{\mathrm{rng}} + a_{k\tau}^{\mathrm{g}} = D_{k\tau}^{\mathrm{g}} \qquad (9a)
\end{aligned}
$$

constraints (9a) state that for each node and period, the imported NG (i.e., injection), flow to other NG nodes, flow to power nodes, flow from and to storage nodes, satisfied load by RNG should, and unsatisfied NG load add up to demand. Unlike power flow, the flow in pipelines are modeled unidirectional as it is typical for most long-distance transmission pipelines involving booster compressor stations [**?**]. We are ignoring electricity consumption associated with booster compression stations along the NG pipeline network. Note that there is no load shedding in NG system as we assume RNG availability for any quantity.

Representative Days:

$$
f_{kn\tau_1}^{\mathrm{ge}} = f_{kn\tau_2}^{\mathrm{ge}} \qquad \tau_1, \tau_2 \in \Omega_\tau, \tau \in \mathfrak{R} \qquad (10a)
$$

The constraint (10a) captures the impact of representative days on the gas system. It ensures that gas consumption by the power system for all the days in the same cluster is the same.

Gas and RNG Supply Constraints: For every $k \in \mathcal{N}^{\mathrm{g}}, \tau \in \mathcal{T}^{\mathrm{g}}$

$$
\begin{aligned}
& L_k^{\mathrm{inj}} \leq g_{k\tau} \leq U_k^{\mathrm{inj}} && (11a) \\
& a_{k\tau}^{\mathrm{rng}} + a_{k\tau}^{\mathrm{g}} \leq D_{k\tau}^{\mathrm{g}} && (11b) \\
& a_{k\tau}^{\mathrm{rng}} \leq U_k^{\mathrm{inj}} M && (11c)
\end{aligned}
$$

import limits are imposed in constraints (11a). The consumption of RNG plus the load shedding is limit by constraints (11b) to the NG load. The alternative fuel RNG can only be imported from injection points as specified by constraints 11c.

Flow Constraints: For every $\ell \in \mathcal{L}^{\mathrm{g}}, \tau \in \mathcal{T}^{\mathrm{g}}, j \in \mathcal{N}^{s}$

$$
\begin{aligned}
& f^{\mathrm{g}}_{\ell\tau} \le I^{\mathrm{pipe}}_{\ell} && \text{if } \mathcal{I}^{\mathrm{pipe}}_{\ell} = 1 && (12a)\\
& f^{\mathrm{g}}_{\ell\tau} \le U^{\mathrm{pipe}}_{\ell} z^{\mathrm{g}}_{\ell} && \text{if } \mathcal{I}^{\mathrm{pipe}}_{\ell} = 0 && (12b)\\
& \sum_{k \in \mathcal{N}^{\mathrm{g}}: j \in \mathcal{A}^{s}_{k}} f^{\mathrm{gl}}_{kj\tau} = s^{\mathrm{liq}}_{j\tau} && && (12c)\\
& \sum_{k \in \mathcal{N}^{\mathrm{g}}: j \in \mathcal{A}^{s}_{k}} f^{\mathrm{vg}}_{jk\tau} = s^{\mathrm{vpr}}_{j\tau} && && (12d)
\end{aligned}
$$

constraints (12a) and (12b) limit the flow between NG nodes for existing and candidate pipelines, respectively. The flow to liquefaction facilities is calculated in constraints (12c). Similarly, the flow out of vaporization facilities is modeled via constraints (12d). note that because these SVL nodes are often connected with the NG nodes with truck transport, that is not modeled here, we do not consider capacity constraint related to the flows to and from the SVL nodes.

Storage Constraints: For every $j \in \mathcal{N}^{s}, \tau \in \mathcal{T}^{\mathrm{g}}$

$$
\begin{aligned}
& s^{\mathrm{gStr}}_{j\tau} = (1-\beta) s^{\mathrm{gStr}}_{j,\tau-1} + \gamma^{\mathrm{liqCh}}_{j} s^{\mathrm{liq}}_{j\tau} - \frac{s^{\mathrm{vpr}}_{j\tau}}{\gamma^{\mathrm{vprDis}}_{j}} && (13a)\\
& s^{\mathrm{vpr}}_{j\tau} \le I^{\mathrm{vpr}}_{j} + x^{\mathrm{vpr}}_{j} && (13b)\\
& s^{\mathrm{gStr}}_{j\tau} \le I^{\mathrm{gStr}}_{j} + x^{\mathrm{gStr}}_{j} && (13c)
\end{aligned}
$$

constraints (13a) ensure the storage balance. Constraints (13b) and (13c) limit the capacity of vaporization and storage tanks to their initial capacity plus the extended capacity, respectively.

1.3 Coupling Constraints

The following constraints are coupling constraints that relate decisions of the two systems.

$$
\begin{aligned}
& \sum_{k \in \mathcal{A}^{\mathrm{e}}_{n}} f^{\mathrm{ge}}_{kn\tau} = \sum_{t \in \mathfrak{T}^{\mathrm{e}}_{\tau}} \sum_{i \in \mathcal{G}} h_i p_{nti} && n \in \mathcal{N}^{\mathrm{e}}, \tau \in \mathfrak{R} && (14a)\\
& \mathcal{E}^{\mathrm{e}} = \sum_{n \in \mathcal{N}^{\mathrm{e}}} \sum_{t \in \mathcal{T}^{e}} \sum_{i \in \mathcal{G}} w_t (1-\eta_i) \eta^{\mathrm{g}} h_i p_{nti} && &&\\
& \mathcal{E}^{\mathrm{g}} = \sum_{k \in \mathcal{N}^{g}} \sum_{\tau \in \mathcal{T}^{g}} \eta^{\mathrm{g}} (D^{\mathrm{g}}_{k\tau} - a^{\mathrm{rng}}_{k\tau} - a^{\mathrm{g}}_{k\tau}) && &&\\
& \mathcal{E}^{\mathrm{e}} + \mathcal{E}^{\mathrm{g}} \le (1-\zeta)(U^{\mathrm{e}}_{\mathrm{emis}} + U^{\mathrm{g}}_{\mathrm{emis}}) && && (14b)
\end{aligned}
$$

The first coupling constraints (14a) captures the flow of NG to the power network for each node and at each time period. The variable $\mathcal{E}^{\mathrm{e}}$ accounts for the emission due to the consumption of NG in the power system. The variable $\mathcal{E}^{\mathrm{g}}$ computes the emission from NG system by subtracting the demand from RNG consumption. The second coupling constraint (14b) ensures that the net CO_2 emissions associated with electric-NG system is below the specified threshold value, which is defined based on reduction relative to some baseline emissions.Since the model cannot track whether RNG is used to meet non-power NG demand or for power generation, the constraint (14b) first computes gross emissions from all NG use presuming it is all fossil and then subtracts emissions benefits from using RNG. Here we treat RNG as a carbon-neutral fuel source, and thus the combustion emissions associated with its end-use are equal to the emissions captured during its production. Recent life cycle analysis studies suggest that depending on the feedstock RNG could have negative to slightly positive life cycle GHG emissions [**?**].

The first term is the emission due to non-generational NG consumption (i.e., NG consumption in the NG system such as space heating, industry use, and transportation) and the second term captures the emission from gas-fired power plants. Alternatively, the emission constraints can only be applied to the power system as in [4] or separately applied to each system as in [**?**].

References

[1] Aron Brenner, Rahman Khorramfar, Dharik Mallapragada, and Saurabh Amin. Graph representation learning for energy demand data: Application to joint energy system planning under emissions constraints. *arXiv preprint arXiv:2209.12035*, 2022.

[2] Wesley J Cole, Danny Greer, Paul Denholm, A Will Frazier, Scott Machen, Trieu Mai, Nina Vincent, and Samuel F Baldwin. Quantifying the challenge of reaching a 100% renewable energy power system for the united states. *Joule*, 5(7):1732–1748, 2021.

[3] Can Li, Antonio J Conejo, John D Siirola, and Ignacio E Grossmann. On representative day selection for capacity expansion planning of power systems under extreme operating conditions. *International Journal of Electrical Power & Energy Systems*, 137:107697, 2022.

[4] Nestor A Sepulveda, Jesse D Jenkins, Aurora Edington, Dharik S Mallapragada, and Richard K Lester. The design space for long-duration energy storage in decarbonized power systems. *Nature Energy*, 6(5):506–516, 2021.