Future Energy Systems Center Pre-publication Paper

This is a pre-publication paper under review by a journal. It can only be shared within Future Energy Systems Center sponsor companies and cannot be made public. When sharing within your sponsor company, please state that this paper cannot be shared outside of the company.

MITei
MIT Energy Initiative

arXiv:2308.16814v2 [eess.SY] 28 Feb 2024

Cost-effective Planning of Decarbonized Power-Gas Infrastructure to Meet the Challenges of Heating Electrification

Rahman Khorramfar[1,2†], Morgan Santoni-Colvin[1†], Saurabh Amin[2,3], Leslie K. Norford[4], Audun Botterud[2] and Dharik Mallapragada[1*]

[1]MIT Energy Initiative, Massachusetts Institute of Technology, Cambridge, MA, 02139.
[2]Laboratory for Information & Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, 02139.
[3]Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139.
[4]Department of Architecture, Massachusetts Institute of Technology, Cambridge, MA, 02139.

*Corresponding author(s). E-mail(s): dharik@mit.edu;
Contributing authors: khorram@mit.edu; msantoni@mit.edu; amins@mit.edu; lnorford@mit.edu; audunb@mit.edu;
†Equal contributions with names in alphabetical order

Abstract

Building heat electrification is central to economy-wide decarbonization efforts and directly affects energy infrastructure planning through increasing electricity demand and reducing the building sector's use of gas infrastructure that also serves the power sector. Here, we develop a modeling framework to quantify end-use demand for electricity and gas in the buildings sector under various electrification pathways and evaluate their impact on co-optimized bulk power-gas infrastructure investments and operations under deep decarbonization scenarios. Applying the framework to study the U.S. New England region in 2050 across 20 weather scenarios, we find high electrification of the residential sector can increase sectoral peak and total electricity demands by up to 56-158% and 41-59% respectively relative to business-as-usual

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

projections. Employing demand-side measures like building envelope improvements under high electrification, however, can reduce the magnitude and weather sensitivity of peak load as well as induce overall efficiency gains, reducing the combined residential sector energy demand for power and gas by 28-30% relative to the present day. Notably, a combination of high electrification and envelope improvements yields the lowest bulk power-gas system cost outcomes. We also find that inter-annual weather-driven variations in demand result in up to 33% variation in optimal power sector investments, which highlights the importance of capturing weather sensitivity of demand for planning purposes.

Keywords: Heating electrification, Joint power-gas planning, inter-annual weather variation, decarbonization

1 Introduction

Economy-wide decarbonization by mid-century requires accelerated CO_2 emissions reduction efforts across all sectors of the economy. Yet, to date, the building sector has seen limited progress in decarbonization efforts relative to other sectors. For example, in the U.S. between 2005 and 2021, power sector CO_2 emissions declined by 34% while the total emissions in the building sector declined by only 18% [1]. Natural gas (NG) remains a significant contributor to the sector's emissions. As of 2021, 80% of on-site fossil fuel consumption in the residential sector, which accounted for 5% of total U.S. GHG emissions, is from NG, with heating being the dominant use case [2]. This share of emissions can be higher in cold climate regions with more NG heating; for example, in New England, residential NG usage accounted for 8% [3] of economy-wide emissions in 2021. The electrification of space heating via efficient air-source heat pumps (ASHPs), combined with a low-carbon power sector, is one of the most widely recognized and viable pathways to significantly reduce buildings sector emissions [4, 5]. Accordingly, federal and state policies increasingly aim to facilitate the deployment of heat pumps [6–8].

The electrification of space heating, however, introduces new challenges for long-term planning and operations of electricity and gas infrastructure under emissions constraints. First, electrification is anticipated to increase the magnitude and alter the pattern of electricity consumption by buildings while displacing NG consumption [9–11]. For example, prior studies of heat electrification suggest that cold climates like the U.S. Northeast will experience a 2.5-3.9 times increase in state-level peak electricity demands for the building sector, with the peak occurring in the winter rather than the present summer-peaking conditions [12]. Importantly, the projected impacts of heating electrification on final energy demand for electricity and NG are strongly dependent on the building stock as well as technology deployment assumptions including the *size* of the heat pumps. For example, under a *whole-home electrification* approach, ASHPs are sized to serve heating needs throughout

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

the year, including during the coldest periods. In contrast, a *hybrid electrification* approach would use smaller ASHPs coupled with backup heating fuel, where the latter is used during especially cold periods to reduce peak electricity consumption and requisite power infrastructure investment but may continue to need the gas infrastructure [12]. While these strategies have each been studied individually [10–13], heterogeneous consumer behavior and policy will result in the simultaneous adoption of multiple sizing strategies across the housing stock. Furthermore, the bulk of the literature has overlooked the role of building envelope improvements in reducing the demand impacts of either electrification strategy.

A second challenge is that the altered electricity and gas consumption patterns due to heat electrification will also impact the utilization of bulk gas infrastructure by reducing its relevance for heating and potentially changing its role for the grid [14, 15]. To date, only a few studies have considered coordinated planning of regional power-gas infrastructure under high electrification and emissions constraints [14, 15]. Third, heating electrification uniquely increases the sensitivity of electricity demand to weather [16–18], which, along with the weather-dependence of variable renewable energy (VRE) electricity supply, complicates decarbonized grid infrastructure planning [16, 19–21]. For example, resource adequacy assessments of decarbonized grids with high levels of heat electrification need to consider the extent to which cold-weather peak events correlate with VRE production, which requires considering inter-annual weather variations.

Several previous studies have used planning models to evaluate power sector decarbonization pathways under scenarios of high heating electrification [9, 22, 23]. However, these grid-centric studies overlook the resulting impacts of heating electrification on the gas infrastructure and the potential cost impacts from the reduced use of gas for heating. Moreover, the impact of inter-annual weather variations and end-use technology adoption on final energy demand from buildings under high heating electrification and its effects on bulk energy infrastructure needs has not been extensively analyzed. This paper presents a quantitative framework to assess the bulk power and gas infrastructure implications of building heat electrification under deep decarbonization scenarios while considering inter-annual weather variations and a range of technological interventions on both demand and supply side. Our key contributions include: i) conducting a spatio-temporally resolved evaluation of electrification's impact on the residential demand for power and gas under several end-use adoption scenarios; ii) modeling the power and gas infrastructure interactions under emissions constraints, including quantifying the system impacts and costs of shifting demand from gas to electricity in the building sector; and iii) considering weather-induced temporal variations in demand and supply for both gas and electricity and implications for power system balancing.

Our work also includes the following key methodological contributions. First, we develop a bottom-up model to project electricity and gas demand

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

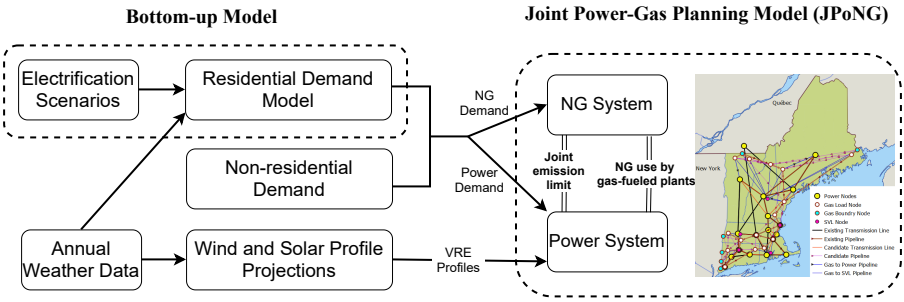


Fig. 1: Overview of the modeling framework to evaluate the impact of residential heating electrification on energy infrastructure planning outcomes. Residential demand is explicitly modeled in the study to consider impact of various demand-side technological interventions. Nonresidential power and gas demand is held constant across all scenarios evaluated as per the projections for high electrification scenarios available from another study [24].

for residential buildings under several electrification pathways and weather scenarios that explicitly consider building stock heterogeneity, heat pump sizing, and inter-annual weather variations in the analysis. Second, we formulate an optimization model for joint planning of electricity and gas infrastructure that is capable of exploring strategies for deep decarbonization of the joint system under various scenarios of electrification output from the bottom-up modeling [14]. The model accounts for weather-induced temporal variations in demand and supply as well as the availability of alternative technologies including renewable generators (VRE and hydropower), short-duration battery storage (Li-ion), low-carbon fuels (LCF) that can substitute for NG, gas generation with and without carbon capture and storage (CCS), demand-side flexibility and transmission expansion (see Fig. 1).

Applying our models to the case study of the New England region in the U.S. in 2050 reveals the following major observations. **First**, the electrification of heating in approximately 60-80% of the residential housing stock under different sizing strategies results in between -8 to 158% increases in peak sectoral electricity demand at the regional level relative to business-as-usual scenario. The aggregate peak sectoral demand change masks significant sub-regional variations - for instance, less populated regions (e.g. Vermont, Maine and rural New Hampshire) experience the greatest relative increases in peak load. **Second**, high electrification (i.e. electrifying ∼80% of homes with an emphasis on whole-home electrification) coupled with envelope improvements results in the lowest combined residential power and gas demand that is up to 28-30% lower than values in the year 2021 and reduces inter-annual variation in peak electricity demand. **Third**, combining high electrification of the residential sector with building envelope improvements incurs the lowest cost of bulk energy infrastructure (i.e. not including housing stock retrofit or distribution infrastructure costs) needed to meet sectoral emissions reduction of 80-95% relative to 1990 levels.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Fourth, while cost-optimal power sector decarbonization strategies rely heavily on VRE supply (mostly wind), they are also accompanied by transmission expansion (10-41% higher than the present), short-duration storage and low-carbon firm power resources such as gas-based generation with and without CCS, fueled by a combination of LCF and fossil gas (referred to as NG henceforth). Notably, across the scenarios, the levels of total NG consumption are less sensitive to the extent of electrification than the level of power-gas system decarbonization, implying substitution effects between NG use in buildings and power. At the same time, the role of LCF is found to grow significantly under certain electrification and decarbonization scenarios. **Finally**, we isolate and quantify the bulk energy infrastructure impact of inter-annual weather-driven variations in energy demand under high electrification. In particular, variation in demand across 20 weather years is found to induce variation in cost-optimal investments in certain generation technologies up to 33%, suggesting that planners must account for weather-induced demand variation in planning future decarbonized energy systems.

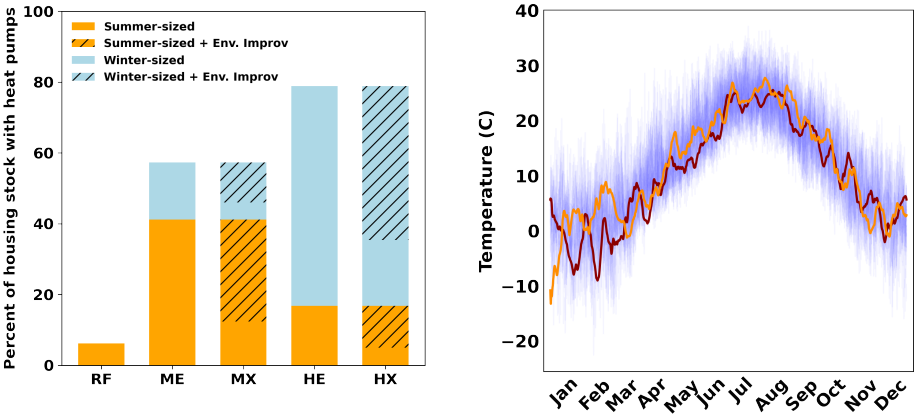
2 Demand impacts of heating electrification

We consider five residential electrification scenarios for 2050, each with different mixes of heat pump adoption, sizing, and envelope improvements that include basic retrofits to improve the thermal efficiency of the building exterior (see SI 12.3.2 and 12.1.3) as shown in Fig. 2a. The first four scenarios include medium and high electrification levels (ME, HE) and variations of them with envelope improvements (MX, HX) that are parameterized to mimic 2050 electrification policy goals in the region. We also evaluate a reference scenario with lower electrification levels to reflect business-as-usual adoption rates (RF). We evaluate power-gas demand for these scenarios under 20 different projections of annual weather data that are based on historical weather patterns and adjusted for the effects of climate change through 2050 (Section 5.1), which we refer to as *weather years*. As an example, Fig. 2b depicts the diversity of the temperature profiles for the full set of weather years for the Boston area. We also simulate so-called "present-day" hourly demand estimates using our bottom-up method in order to compare the 2050 demand changes to current system conditions and to benchmark the accuracy of our modeling approach (see SI 12.3.4).

2.1 Substitution of NG demand by electricity demand

With increasing amounts of heating electrification in the housing stock, Fig. 3a shows that annual residential NG consumption declines while electricity consumption increases. For scenarios HE and ME, residential electricity demand increases over the present day by an average of 33.2 TWh (63%) and 16.8 TWh (32%) across all weather years, respectively. The impact of these electrification scenarios on site-level gas demand (excluding power generation) is a reduction of 55.4 TWh (71%) and 29.2 TWh (38%). Due to the higher efficiency

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.



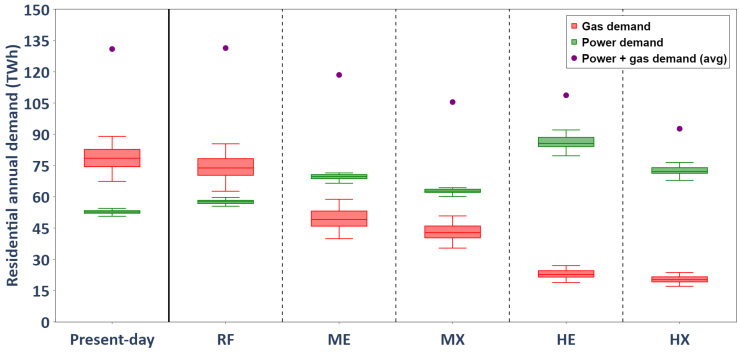
(a) (b)

Fig. 2: **(a)** Scenarios for electrification of residential building heating for 2050. "Winter-sized" homes have large whole-home heat pumps that are sized to provide heat through the winter and are used as the only source of heating and cooling. Smaller "Summer-sized" heat pumps are primarily sized to meet air conditioning needs in the summer, but also provide some heating in the winter, supplemented by a backup heating system. In our analysis, "Summer-sized" systems are almost all also hybrid systems because the existing heating systems in New England homes are typically gas or oil. In the MX and HX scenarios, we assume 70% of all homes given heat pumps also receive envelope improvements. See further details in SI 12.3.2. **(b)** Temperature profiles in Suffolk County, MA for the 20 weather years used in the analysis. Light blue lines in the background illustrate hourly temperature values for all 20 years overlaid on top of one another to illustrate the weather variations. Orange and bold red lines highlight weekly rolling averages of temperatures for two different years drawn from the dataset. A more detailed depiction of the 20 years of temperature data can be found in Fig. S30.

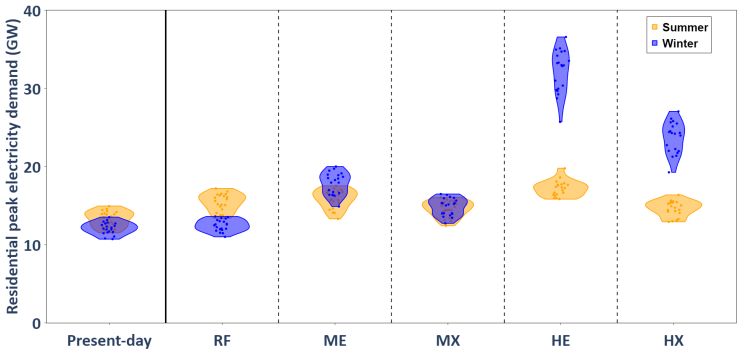
of heat pumps compared to the existing heating stock, the total combined end-use energy demand for power and gas also declines relative to the present day with increasing electrification, by 16-18% and 28-30% for the HE and HX scenarios respectively. At increasing levels of electrification, inter-annual variation in demand decreases for gas but increases for electricity because greater amounts of the heating load are being met by electricity. Deploying envelope improvements in parallel with heat pumps reduces the increase in electricity consumption; consumption of electricity is reduced by an average of 16% between scenarios HE and HX, and by 10% between scenarios ME and MX. Envelope improvements also induce average gas demand savings of 11% and 13%, respectively.

Since our modeled electrification scenarios include some hybrid heating systems that leverage the existing gas heating system as backup, their operation alone is responsible for 2.8-8.8 TWh of gas consumption across the weather years in the high electrification scenarios (HE, HX) and 6.9-21.6 TWh for the medium electrification scenarios (ME, MX).

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.



(a)



(b)

Fig. 3: **(a)** Annual electricity and gas demands in residential sector under different electrification scenarios. Each box plot contains demand data for the 20 weather years. HE, HX, ME, MX, and RF are all 2050 scenarios. The box edges correspond to interquartile range (IQR) and the whiskers indicate the most extreme points within the range Q1 - 1.5*IQR to Q3 + 1.5*IQR. **(b)** Summer and winter peak demands for the residential sector under different electrification scenarios. Each violin contains demand data simulated for 20 weather years. The width of plots increases where the density of data is higher, and its height represents sensitivity toward the weather variation. For New England, our "present-day" model and data workflow result in overestimation of residential gas consumption as compared to published outputs of ResStock runs by NREL [25] as well as historical annual values available from EIA (shown in Figure S3). As discussed in SI 7.3, the difference between our present-day model results and EIA data is likely attributable to a combination of a) differences in weather data used in our analysis (see Fig. S4), b) the limited number of building archetypes we consider to maintain computational tractability, and c) error within the ResStock model. The analysis presented in 7.3 suggests that the error in the present-day model does not as strongly impact the demand results from our future electrification scenarios.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

2.2 Changes in residential electric peak demand

Similar to other studies on heat electrification, we find that peak electricity demands, which are generally an important metric for grid planning purposes, increase with electrification and become more weather-dependent, as highlighted in Fig. 3b. The weather dependence of peak demand is illustrated by the fact that the maximum and minimum annual peak demands across the 20 weather years differ by 3.0 GW for the present-day and 10.9 GW for HE. Because the peak demand generally coincides with extreme cold weather, this result suggests that higher levels of electrification will magnify the adverse peak demand impacts during periods of extreme cold. Importantly, increases in peak electricity demand at high electrification are greater than the increases in annual demand - for example, peak electricity demand and annual electricity demand for the HE scenario, which has the highest demand, are 56-158% and 41-59% greater than the RF scenario across 20 weather years, respectively.

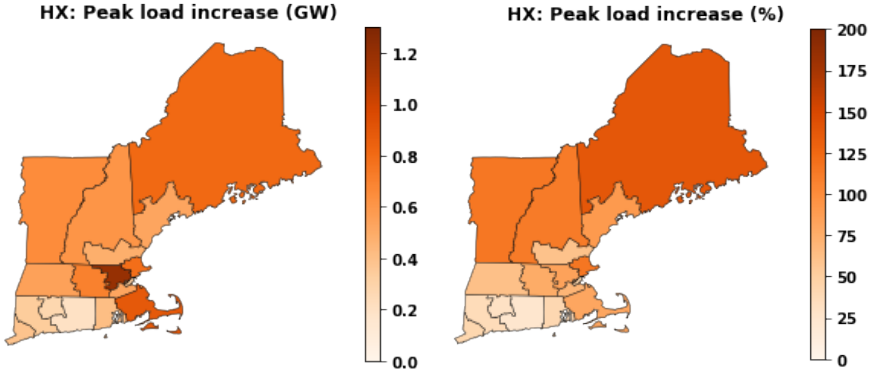
Demand-side measures like envelope improvements can partly mitigate increases in peak demand with electrification as well as reduce its sensitivity to weather variations. Envelope improvements provide reductions in peak demand by an average of 8.5 GW for HX compared to HE (resulting in a peak 17% to 92% higher than RF) and 2.5 GW for MX compared to ME (resulting in a peak between 8% lower and 21% higher than RF). They also reduce the duration of peak electricity demand events. These events are characterized as the consecutive hours immediately preceding and following the seasonal peak, during which demand exceeds 75% of the demand recorded in the peak hour. The longest peak demand event can last up to 60 hours in the case of the HE scenario but is reduced to 33 hours or lower in the HX scenario (Fig. S1). The supply-side results shown in the next section indicate that the reduction in the frequency, duration and magnitude of peak demand enabled by envelope improvements leads to reduced power capacity investment and lower total power-gas system costs under deep decarbonization scenarios.

Fig. 3b highlights that high electrification of the residential sector alone results in sectoral peak demands that are comparable to the present-day New England system-wide peak demands of 24.4-26.0 GW [26]. While the residential sector's average summer peak in the present-day scenario currently exceeds the winter peak by an average of 1.2 GW, the residential winter peak exceeds summer peak by an average of 0.3, 2.0, 9.0, and 15.0 GW under the MX, ME, HX, and HE scenarios, respectively. New England's system-wide summer peak currently exceeds the winter peak by an average of 5.4 GW [26]. Thus, depending on peak timing in other sectors, HE and HX can drive New England into a winter-peaking system due to the effects of residential electrification alone.

Installing electric resistance heat as backup in lieu of hybrid systems with fuel backup can further reduce fuel consumption without needing a larger heat pump. However, using electric resistance backup comes at the expense of increasing electricity consumption proportionately. Figure S2 shows that using electric resistance backup instead of existing backup further increases peak loads by up to 184% across the various electrification scenarios. This

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

increase in peak electricity demand, which is particularly drastic for the ME and MX scenarios for their higher deployment of smaller heat pumps, may not be economical from a supply-side perspective as it may reduce utilization of dispatchable electricity generation.



(a) (b)

Fig. 4: Average residential peak load increase across the 20 weather years for scenario HX above present-day for 17 zones in New England. A summary of the zone geography is presented in SI 9.1.1. Numerical results shown in Table S3. **(a)** Results in GW and **(b)** on percent basis.

While regionally aggregate peak demand is an important metric for transmission planning, the spatial distribution of peak loads is important for distribution network planning. Fig. 4 shows the average load, both in absolute and percent terms, for scenario HX across the weather years for each load zone. We find that while absolute load increases, as shown in Fig. 4a, are distributed diversely across the region, Fig. 4b reveals that percent load increases relative to the present day are greater in the more rural, colder regions. For example, the state of Vermont incurs an average residential peak demand growth of 115% under the HX scenario with a population of 647,064 [27], while Suffolk County, where Boston is located, has a much lower peak growth of 45% with a population of 766,381 [28]. A key potential implication is that the demand increase in rural areas could require deeper distribution network buildout relative to the population of ratepayers who will likely incur the cost through retail tariffs. Conversely, the more densely populated southern regions of New England may not need as dramatic network upgrades, the costs of which can also be spread across the larger population.

3 Supply-side impacts of heating electrification

We evaluate the impact of different electrification scenarios and inter-annual weather variation on bulk power-gas system outcomes using a modified version of the *JPoNG* model [14]. The model co-optimizes bulk power and gas

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

infrastructure to minimize the sum of their annualized investment and operational costs while adhering to constraints related to technology operation, resource availability, and policy, namely emissions constraints and resource adequacy requirements (see details in Section 5.2 and SI 8). We evaluate model outcomes for an emissions goal of 80% sectoral emissions reduction relative to 1990 levels, consistent with regional policy goals for 2050, as well as a more ambitious policy of 95% emissions reduction [29, 30]. Key model decision variables relate to investment and operation of VRE, short-duration battery storage, power, and gas transmission as well as gas generation with and without CCS. For the latter, we model two gas fuel options: NG imports as well as well as imports of LCF. The latter is presumed to be carbon-neutral and compatible with the gas infrastructure but is available at cost premium relative to NG according to a supply curve (10-50 $/MMBtu vs. $5.45/MMBtu - see Table 1 in Section 5). LCF is meant to serve as a proxy for synthetic or biogenic methane that has been considered by other studies modeling regional and national decarbonization [30–32].

3.1 Cost-effective resource portfolios

To illustrate the least-cost resource portfolios across various electrification and emissions scenarios, Fig. 5 shows the model's outcomes for an exemplar weather year that has the highest combined energy demand for power and gas across the 20 years we evaluate. We include results across all weather years in SI 7.11 (see Figure S11).

Across electrification and decarbonization scenarios, Fig. 5a-b shows that VRE resources make up the majority of capacity and generation. Despite being deployed in smaller amounts than solar in terms of capacity, onshore and offshore wind dominate the generation owing to their higher availability (capacity factor of 50% for offshore wind vs. 19% for solar PV in the exemplar weather year), especially during periods of increased demand in the winter months. For the same reason, as electrification increases, offshore wind is the primary resource used to meet the incremental demand - for example, offshore wind's share of total generation increases from 24% to 34% between the RF and HE scenarios under 95% decarbonization goal. Across all scenarios evaluated in Fig. 5, solar and onshore wind are built to their potential capacity limit in the region [29], implying that increasing land availability for these resources could further increase their deployment in lieu of other resources and reduce system costs.

All electrification scenarios maintain approximately 12.1-15.8 GW of short-duration battery storage (Li-ion) with a rated duration of 3.5-4.3 hours, highlighting the fact that battery storage is required to compensate for the variable nature of VRE, particularly solar. The medium and high electrification scenarios drive transmission expansion by 18-30% compared to current network capacity across the emissions scenarios (9.1.3), as reflected by the

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

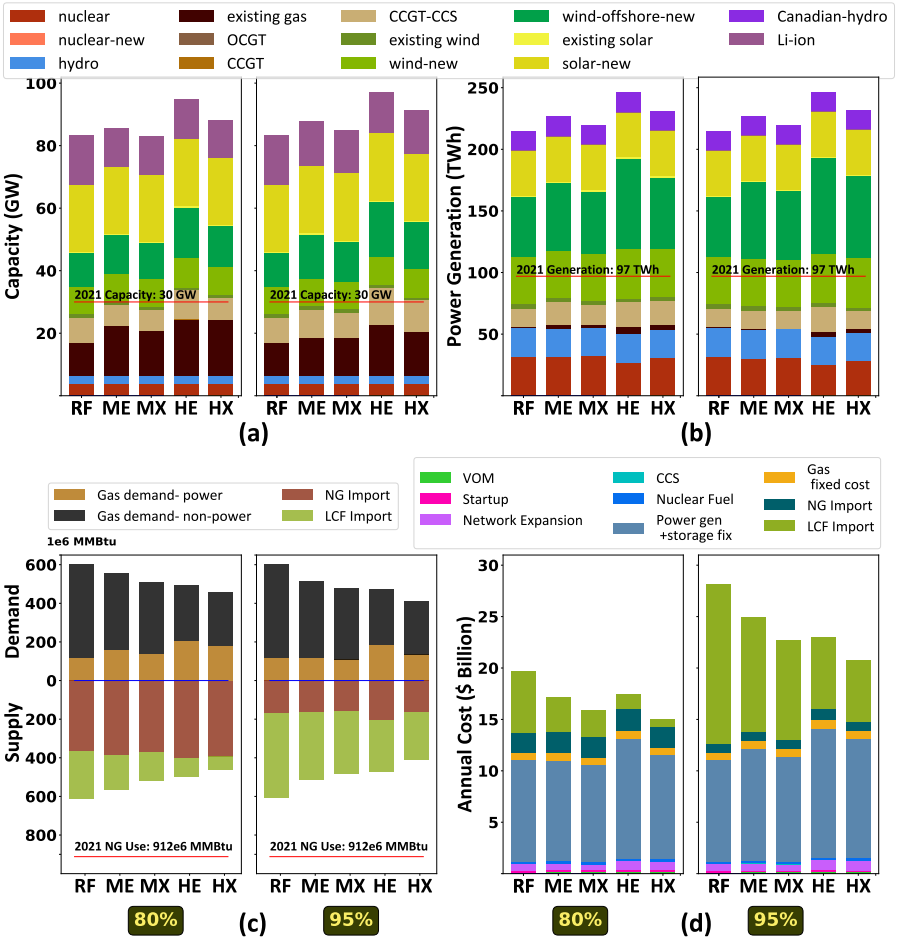


Fig. 5: System outcomes for the power and gas systems for the exemplar weather year. (**a**) Power capacity, (**b**) power generation. 'OCGT' and 'CCGT' are open and combined cycle generators, respectively. 'CCGT-CCS' is a 'CCGT' plant with carbon capture and storage technology. (**c**) gas consumption (top) and supply (bottom), and (**d**) annual power-gas system costs for exemplar weather year under different electrification scenarios, as described in Fig. 3, and decarbonization targets. The investment costs are annualized. In the upper legend, "Li-ion" is short-duration battery storage. In the lower-right legend, "VOM" and "FOM" are variable and fixed operating and maintenance costs, respectively. "CCS" is the cost of establishing carbon capture and storage infrastructure. 'Power gen + storage fix' is the investment and FOM costs for storage and power plants. 'Network Expansion' is the cost of establishing new transmission lines.

relative cost contribution of network expansion in Fig. 5d. The increased network capacity allows more investment in regions with higher VRE potential (see Fig. S5).

As can be seen by comparing HE to ME in Fig. 5c, increasing electrification also reduces overall gas consumption by reducing its use in the building

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

sector which more than offsets any increased use of gas in the power sector. This outcome is likely partially driven by the aforementioned efficiency benefits of heat pumps. Moreover, the LCF consumption for HE and HX scenarios varies between 70e6 to 11e6 MMBtu under 80%, and between 217e6 to 277e6 MMBtu under 95% decarbonization target; this is almost within the production capacity of the region for the biogenic LCF (see Table 1 in Section 5) and highlights the need for high electrification. Implementation of envelope improvements coupled with electrification on the supply side generally reduces power capacity investments, which is most notable for high electrification scenarios (Fig. 5a), as well as reduces gas consumption across power and building sectors (Fig. 5c).

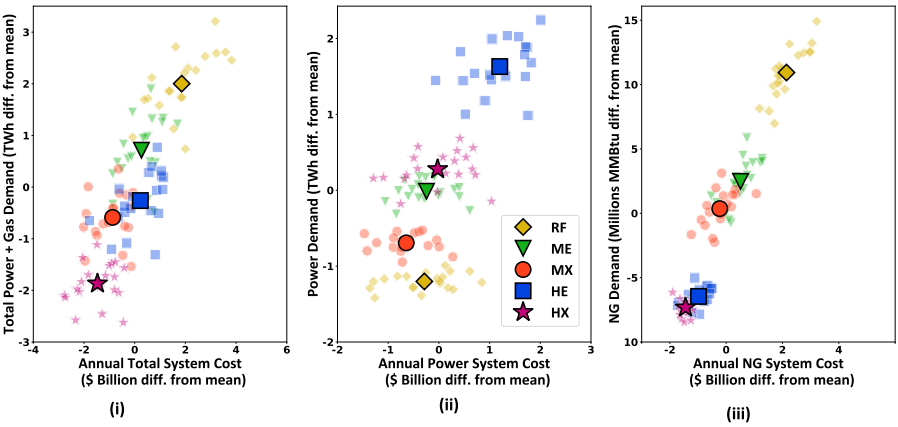
The comparison between Fig. 5a and 5b also reveals that the system retains most of the existing unabated gas-fired power plants but these plants are utilized in very limited amounts (with capacity factors of 0.6-3.5%) and are primarily used to meet peak demand during extreme cold-weather events (as shown in SI 7.7). Fig. 5a also shows that CCS-based plants are highly utilized, with capacity factors ranging from 17% to 31%. At 18.5-28.3 GW, the overall capacity of the fleet of gas power plants is roughly 4-59% larger than the existing capacity but makes up only 7.2-11.4% of overall generation as compared to 52% for New England in 2022 [33]. Motivated by the low regional CO_2 storage availability and lack of regional policy interest in CCS, we evaluated a sensitivity case without CCS for the electrification and decarbonization scenarios. The results (Fig. S7) point to an increased reliance on offshore wind and LCF that is used via unabated gas generation to balance the system in the absence of CCS, which can be done at similar or slightly increased cost (see Fig. S8).

Fig. 5c reveals that a more stringent decarbonization goal results in higher LCF deployment (217e6 to 277e6 MMBtu), nearly doubling across all electrification scenarios for the 95% emissions reduction target, and also making a large impact on the combined power and NG system cost (see Fig. 5d). The expensive LCF supply comprises a significant fraction of overall cost, ranging from on average 16% in the 80% decarbonization case to 40% under the 95% decarbonization target. This results in a system-wide cost increase of on average 40% across the electrification scenarios when moving from an 80% emissions goal to 95%.

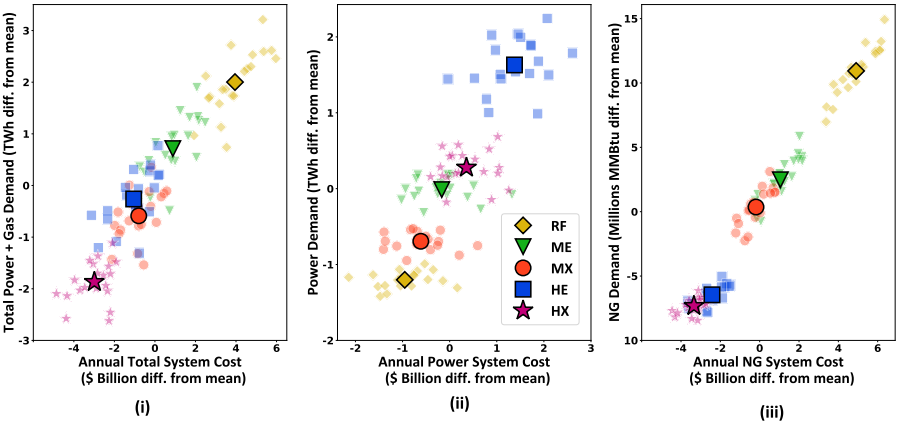
3.2 Impact of electrification pathways on power-gas system costs

Across the weather years and decarbonization scenarios, Fig. 6 highlights that the average bulk power-gas system costs (shown by the emphasized points) are lowest for high levels of electrification that entail envelope improvements. While a full assessment of the cost of envelope improvement retrofits is outside the scope of this work, a simple net-present-value analysis demonstrates that the direct bulk system savings alone offset roughly 35-57% of the cost of envelope improvements under high electrification (see SI 7.12). This analysis does not account for other co-benefits of envelope improvements (see Discussion).

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.



(a) Decarbonization Target = 80%



(b) Decarbonization Target = 95%

Fig. 6: Total bulk power-gas system cost versus energy consumption for (**a**) 80% decarbonization and (**b**) 95% decarbonization. Subfigures (i) compare the combined system cost with the combined energy demand for power and NG. Subfigures (ii) and (iii) have the same individual comparisons for power and gas with their corresponding demands. Axes are normalized according to their difference with the average value across all points. Translucent points show all 20 weather years for each electrification scenario, while bold points indicate the averages. Note that bulk power-gas system costs do not include demand-side costs of electrification, such as heat pump and envelope improvement retrofits as well as electricity and gas distribution infrastructure costs.

It is instructive to note that the RF scenario, corresponding to the least electrification, has the highest average power-gas system costs across both decarbonization targets in Fig. 6, which reinforces the importance of electrification as a strategy for economy-wide decarbonization; notably, the costs for RF are even higher than the scenarios with more electrification that lack any envelope improvements. On average relative to the ME, MX, HE, and HX

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

scenarios, RF yields 8%, 15%, 9% and 18% higher costs under the 80% decarbonization goal and 11%, 18%, 19%, and 26% higher costs under 95% emissions reduction goal, respectively. While higher levels of electrification can lead to increased costs in the power system as illustrated in Fig. 6a(ii) and 6b(ii), they lead to substantial savings in the NG system (Fig. 6a(iii) and 6b(iii)) that partially or entirely offset power system cost increases. This highlights the importance of joint energy system planning and indicates that evaluating the gas or power systems in isolation from each other can lead to misleading results.

3.3 System planning implications of weather sensitivity of demand

Generally, studies focused on decarbonized grid infrastructure planning that consider inter-annual weather variation have tended to emphasize its impact on supply rather than on demand [21, 34–37]. Among studies looking at demand and supply variations [16, 19], the treatment of weather variation in demand under different degrees of heating electrification and its grid impacts is quite limited. To isolate the specific impact of weather-driven demand variations under different electrification levels highlighted in Fig. 3 for a given year of VRE supply data, we compare JPoNG outcomes between the cases where demand data is assessed across various weather years (i.e. *desynchronized cases*) vs. the case where VRE and demand data inputs correspond to the same weather year (*synchronized case*). In Fig. 7, we compare the capacity outcomes between the desynchronized cases and corresponding synchronized cases for the HE and ME scenarios under the 95% decarbonization target.

As revealed in Fig. 7, the weather sensitivity of demand has a significant individual impact on the optimal investment in different plant types. For the desynchronized cases under the ME scenario, investment in offshore wind assets can vary from -16% to 11% relative to the synchronized case. The optimal amount of thermal capacity varies from -20% to +16%, a striking result when considered in the context of the key role these resources play in firming the grid during the extreme demand events highlighted in this study. Solar capacity is essentially insensitive to the demand variation because it is almost always deployed up to the land availability limit regardless of the weather data used. Relatedly, storage capacity variation is similar to offshore wind capacity, from -13% to +13%, presumably because storage primarily serves to balance VRE generation (see Fig. S6).

Comparing the spread of the distributions for the ME and HE scenarios reveals that the increased weather sensitivity of demand at higher electrification levels passes through to the supply side. This results in greater variation in the optimal deployment of certain asset types relative to lower levels of electrification. Between ME and HE, the standard deviation of the distributions for offshore wind and thermal capacity increase by respectively 2.6% and 2.1% of peak demand. In extreme cases for HE, thermal capacity can vary by up to 33% as a result of the demand variation, highlighting the impact that weather

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

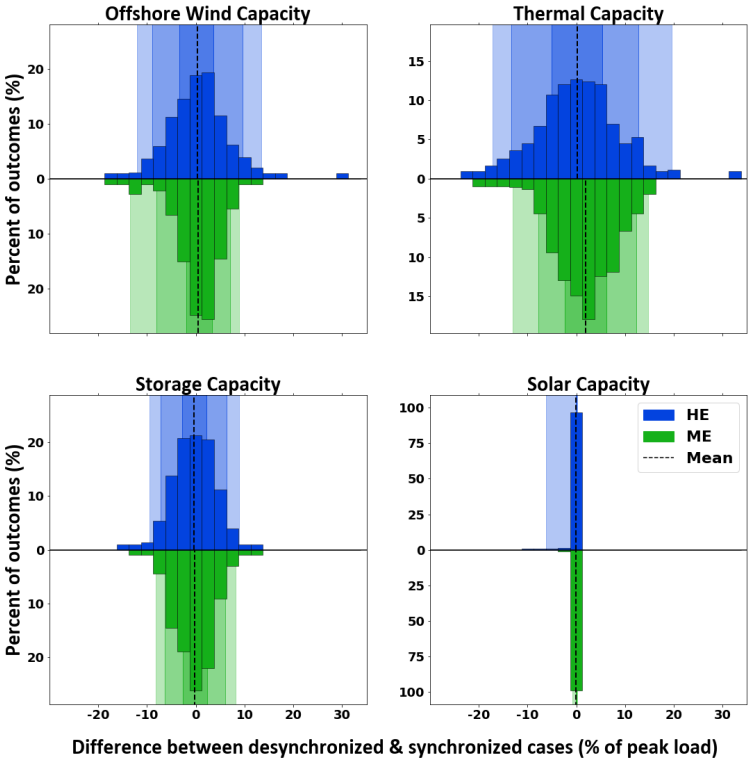


Fig. 7: Deviation in optimal supply side outcomes between the various model runs using different weather years for VRE and demand data (desynchronized cases) and the case where VRE and demand data are from the same weather year (synchronized cases). Results are reported on a normalized basis relative to the peak load of each respective year of demand data. The calculation procedure is as follows: 1) we take a single weather year's VRE profiles and run the model using the demand profiles of each of the 20 weather years. 2) We then calculate the differences in system outcomes between the case for which the VRE and demand data are from the same weather year (*synchronized case*) and the remaining 19 cases for which they are not (*desynchronized cases*). 3) By repeating this process for each of the 20 years of supply data, we build a distribution that captures the sensitivity of the supply outcome to the weather variability of demand. The origin in the x-axis represents the synchronized instance. Moving from outside to inwards, the background shading represents 1st-99th, 5th-95th, and 25th-75th quantiles of the data, respectively. For illustrative purposes, thermal capacity encompasses the sum of capacities of pre-existing gas plants, new OCGT, new CCGT, and new CCGT with CCS.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

extremes can have on cost-effective system outcomes. By contrast, storage and solar are less affected by the level of electrification.

4 Discussion

Our study illustrates the importance of taking an integrated planning approach to heating electrification and applying methods that recognize the importance of weather variations. We show that, without such mitigating measures as envelope improvements, peak electricity demand is likely to significantly increase under deep electrification, necessitating further investment in grid supply resources that may be under-utilized. Regardless, we find that high electrification, even without envelope improvements, can result in reduced bulk power-gas system costs vs. the business-as-usual scenario with modest electrification (where less than 10% of the housing stock is electrified, see scenario RF in Fig. 2a) under both decarbonization scenarios considered. This is in contrast to some suggestions that more aggressive electrification pathways are more expensive due to the capital investment needed on the supply side [38–40]. Although overall bulk power-gas system costs are lower at high levels of electrification, future analysis should evaluate the allocation of these costs, particularly as it pertains to low-income customers who could be stranded on legacy gas infrastructure. In addition, our analysis does not model the cost associated with power-gas distribution as well as end-use technology adoption, which would need to be factored in when considering the cost-benefit analysis of different levels of electrification. For example, hybrid systems are often promoted (and may be favored by consumers) for their potential to reduce the upfront capital costs of ASHP installations and also could reduce the extent of power distribution network upgrades needed to accommodate high electrification levels.

Although they furnish cost savings on the bulk system, envelope improvements also entail retrofit costs at the household level that have not been considered in this analysis. While preliminary calculations suggest that the bulk system cost savings from envelope improvements could partially offset the costs of retrofit implementation (see SI 7.12), there are other co-benefits that are worth noting. These include: avoided distribution infrastructure investment that will be needed to support higher loads, household benefits such as energy bills savings [41], and health benefits [41]. However, these benefits can be undermined by the evidence that traditional energy efficiency measures, including envelope improvements, may not deliver on the full technical potential of their estimated savings [42].

Our results highlight the value of integrated power-gas planning in future decarbonized energy systems, particularly in regions like New England where NG is prevalent as a heating fuel and also currently plays a major role in electricity supply. Our modeling results suggest that electricity demand under decarbonization scenarios will largely be met by VRE including high levels of wind deployment and a complementary mix of solar and storage. However,

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

even under high levels of heating electrification, there may be a need for gas fuel for balancing the power system as well as servicing residual gas demand in the heating stock. These interactions necessitate coordinated planning of the power and gas sectors going forward in order to explicitly evaluate the tradeoffs and explore the full spectrum of potential solutions.

A theme of our findings is that a flexible low-carbon electricity resource and/or low-carbon energy carrier will likely be needed to supply the energy demands of electrification in a cost-effective manner while meeting decarbonization objectives in cold climates. This finding is robust to sensitivities regarding the level of electrification or the availability of CCS. Although the literature recognizes a need for flexible resources more generally [43], this study highlights their role under different levels of heating electrification and interactions with the gas system. While CCS and LCF-based power generation are examples of flexible, low-carbon resources, competing technologies like nuclear-based power generation and long-duration energy storage could also be equally important to consider as highlighted by other studies [14, 43].

In the case of CCS-based power generation and LCF, further scrutiny is needed to understand their role in deeply decarbonized energy systems. First, both technologies face future cost uncertainties due to the lack of their deployment at scale currently. For example, the future supply and cost of LCF are uncertain for reasons such as feedstock supply limitations for biomass and scalability of synthetic methane production [44] [45]. In this context, high electrification serves as an important hedge as it leads to the lowest LCF consumption across the modeled demand scenarios. Moreover, the LCF consumption in these scenarios (HE-HX: 70e6-277e6 MMBtu across both decarbonization targets) is similar to the maximum technical potential of biogenic methane assumed for the region (see SI 3.3). In addition, the GHG emissions benefit of using NG in conjunction with CCS and using LCF with or without CCS could be reduced when factoring methane emissions associated with the supply chains of these fuels. Unsurprisingly, accounting for the supply chain methane emissions for these fuels in the emissions constraints raises the system cost of decarbonization, through increased substitution of NG with less-carbon intensive LCF, as highlighted in SI 7.9. At the same time, the continued reliance on these fuels even when accounting for supply-chain methane emissions, highlights the need for policy mandates to measure and mitigate supply-chain emissions, such as those being undertaken in the Methane Emissions Reductions Program implemented under the Inflation Reduction Act [46].

Our analysis also highlights that the weather-induced demand variability experienced under heating electrification leads to notable variations in the optimal investment outcomes on the supply side, such that key capacity investments can vary by up to 33%. Similar to the focus on the impact of interannual weather variation on VRE generation, planners will have to consider the impact on demand, particularly in a future with high heating electrification. This sensitivity calls for principled planning approaches that account

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

for weather-induced demand variation, or alternatively, the implementation of demand-side solutions to mitigate the variability and enable convergence to a single "best" system plan.

Future work could serve to consider other key dimensions of the modeling problem. Evaluating the hourly impacts of commercial building sector heat electrification would be valuable, but so far has proven to be a difficult bottom-up modeling problem [25]. Future analysis should also consider the role of demand flexibility, which could be instrumental in reducing peak demand and power system costs under electrification, similar to the envelope improvement scenarios modeled here. As an initial investigation into the potential role of demand flexibility under heat electrification, we test the model's sensitivity to flexible transportation loads and find that activating flexibility in this sector could result in significant system-wide cost savings (see SI 7.10). Furthermore, given that our model results call for significant network expansion and power import from neighboring regions (e.g., Canada), the evaluation of resiliency-related issues such as network outages and coordination of supply and demand with neighboring regions during extreme cold weather events will need to be studied further.

5 Methodology

Our framework consists of bottom-up and planning models as depicted in Fig. 1. The bottom-up model projects the sectoral demand of power and gas for the year 2050 under various electrification and weather years. Using the projected values, the planning model provides investment and operational outcomes for each unique instance of weather year and electrification scenario.

5.1 Bottom-up engineering model for energy demand

We develop a bottom-up model to project the electricity and gas demands in the residential sector of the U.S. New England region in the year 2050 using 20 years of historical weather data, climate change projections, and other publicly available data and open-source resources. SI 12 provides a detailed description of the model. Because our focus is on the bulk system, we do not evaluate the demand impacts of electrification on other heating fuels such as fuel oil or propane. Our modeling framework allows us to leverage existing heat-pump adoption rate projections (see 12.3.2) to produce demand profiles for different levels of electrification. We divide New England into 17 zones each representing a population of between 0.6 and 1.6 million people (see SI 9.1.1). The zones serve as the nodes of the power system network and the regions for which demand data is generated. Our process consists of three modeling steps as shown in Fig. 8.

In the first step of the demand modeling, we leverage NREL's ResStock model [25, 53]. Given the housing stock characteristics and weather, ResStock generates a sample of housing archetypes and simulates their hourly energy consumption for key end-uses. We modify ResStock's housing stock

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

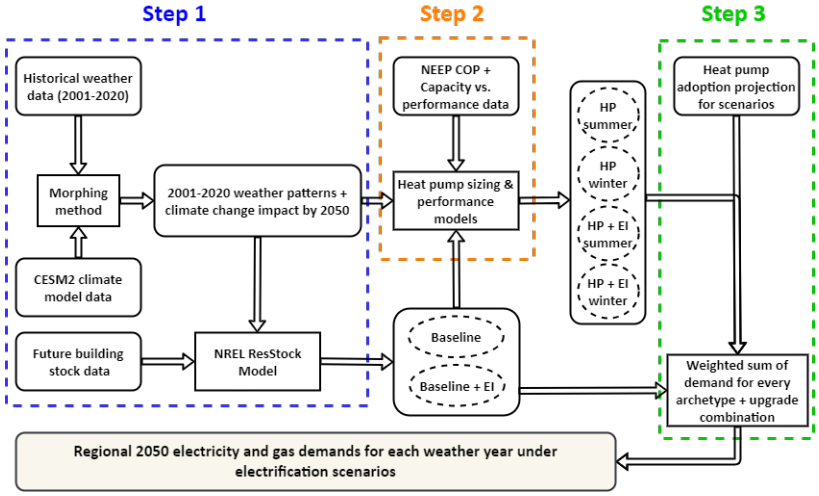


Fig. 8: Diagram of bottom-up residential model. "Summer" and "winter" correspond to the two heat pump sizing methods applied in this study and capture both the whole-home and hybrid approaches to heat pump retrofits (see Figure 2a). "Baseline" corresponds to the remaining homes that do not receive heat pumps. "EI" indicates the subset of homes that receive envelope improvements. See SI 12 for more details. Step 1 (see 12.1) determines heating and gas fuel demands for the original archetypes using future weather and building stock projections. Step 2 determines the electricity and gas demands for the archetypes when they are electrified under different mixes of heat pump sizes and envelope upgrades (see 12.2). Step 3 aggregates the demand profiles of all of the archetypes to determine the hourly regional loads for the residential sector while accounting for the mix of possible heat pump sizes and envelope improvements applied to the stock (see 12.3). Historical weather data is obtained from ERA5 [47] via [48] (see Fig. S29 for more details on rationale for using the ERA5 dataset), morphing method from [49–51], CESM2 climate model data from [52], and future building stock data from [7].

data inputs to resemble the projected housing stock of New England in 2050 (see SI 12.1.1). As the basis for our 20 years of weather inputs, we procure hourly historical weather data for weather years 2001-2020 for 44 locations in New England corresponding to the DOE's typical meteorological year (TMY) locations [54] (see SI 12.4.1). We simulate each archetype with the weather data of the closest location. We account for the climate change impact between the present-day and 2050 by applying the *morphing method*, a common tool in the building energy modeling literature [49–51] to modify the hourly weather data according to monthly climate model projections that characterize the change of key weather metrics over time. More information on the morphing method is included in SI 12.5. ResStock's hourly outputs exhibit a smoothing effect with increasing numbers of archetypes [25]. We run simulations for approximately 400 archetypes in each of the 17 demand zones in order to balance the benefits of the smoothing effect with the computational burden of more simulations. We discuss the implications of sample size on simulation outputs and model error in 7.3 and note that this value has been used previously in similar work [10]. In addition to the baseline stock, we use ResStock to

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

model envelope improvements that may occur in parallel with electrification (see SI 12.1.3).

In the second step, we determine the impacts of electrification for each archetype. We use the heating and cooling demands from ResStock as inputs to a heat pump model. The efficiency of ASHPs to provide cooling or heating decreases as the difference between indoor and outdoor temperatures increases. Maximum heating capacity also degrades at lower temperatures. Therefore, we calculate capacity and efficiency at each hour by fitting a regression model to a large dataset of temperature-dependent heat pump performance. The performance curves are such that the heat pumps have a heating efficiency of 273% (COP of 2.73) at 0 C and 183% at -20 C. Further details about these models are included in SI 12.2.1-12.2.3.

We model the heat pump's energy demand under two possible forms of sizing to capture distinct electrification strategies: 1) whole-home or winter sizing and 2) summer sizing, where the heat pump is supplemented by the existing heating (backup) system, most often in a hybrid configuration because most existing heating systems in the region use fuels rather than electricity. We determine the winter and summer sizes for a given archetype based on heating and cooling demands for a TMY using approximations of heat pump sizing methods taken from industry standards (see SI 12.2.4). For summer-sized systems, we assume the heat pump and backup system operate based on the characteristics and limitations of the existing system in each home as described in SI 12.2.5. End-use demands excluding heating and cooling are taken from the ResStock model. For the electrified homes, we also assume the water heater is converted to a heat pump water heater (see SI 12.1.4).

The third step produces the zone-level demands for each combination of weather year and electrification scenario. To yield the zone-level hourly gas demands for each scenario, we compute a weighted sum of the demand profiles for all of the archetypes in the zone. The weights are equal to the number of homes represented by each archetype and take into account the heterogeneous mix of sizing and envelope improvements that may apply to different portions of the stock represented by each archetype (more detail in SI 12.3) as well as population change through 2050 (SI 12.3.3). Once the residential load is obtained, we use existing sectoral energy demand projections from the literature for 2050 for all non-residential sectors (see SIs 9.1.6 and 9.3.4).

5.2 Joint power-gas planning model

We use an extended version of the JPoNG [14] model to analyze the power-gas system planning in this paper, with a complete mathematical formulation provided in SI 8. The investment decisions include establishing power plants, transmission lines, dedicated pipelines for carbon sequestration, gas pipelines, and power storage technologies. The model also considers decommissioning legacy power assets as well as gas pipelines. Key model constraints include i) power and gas system supply-demand balancing at hourly and daily time resolution, respectively; ii) network constraints for electricity (DC-OPF) and

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

gas (pipe and bubble formulation); iii) capacity reserve margin for electricity system; iv) resource availability limits for VRE, Canadian hydropower imports and LCF; and v) system-wide emissions reduction goals.

The operations of both systems are coupled through two sets of constraints. The first set ensures gas flow (either NG or LCF) to the power system to fuel gas-fired power plants. The second coupling constraint limits the CO_2 emissions caused by the consumption of NG in both systems to a specified target. The model captures an entire year of daily variability of gas demand and hourly operations of the power system over 24 representative days that are selected by the method proposed in [55].

The system is modeled as a semi-isolated energy system with no power or NG export but with the possibility of power, NG, and LCF import. We construct the network for each system using publicly available data detailed in 9. The resulting power and gas networks have 18 (17 nodes inside New England, one node to represent an import point from Quebec, Canada) and 23 nodes, respectively. This level of spatial granularity is sufficient to facilitate detailed analysis and adequately capture the network congestion effects. Overall, the power system has 39 existing and 34 candidate transmission lines and the gas system includes 36 existing and 46 candidate pipelines. SI 9 provides details of assumptions and technical and economic parameters in the power system. Table 1 shows key system parameter values used in this paper. Most of these parameters are sourced from [56–58]. The detailed gas system data and parameters are provided in SI 9. Although we use JPoNG for the case study of New England, the model is broadly applicable to other regions and countries for both local and national level planning. In particular, the data sets we have used are also available for other U.S. regions (e.g. ResStock, ERA5) and thus the workflows developed could in principle be extended to study other regions in the U.S. readily.

The resulting model is a mixed-integer linear program (MILP) which becomes intractable for current commercial solvers as the number of representative days and the size of the power-gas network increase - the typical problem instance considered has 0.6 million continuous variables, 1000 integer variables and 0.75 million constraints. Therefore, we solve each instance approximately with the four-step approach (FSA) described in SI 11. Our solution approach provides a feasible solution whose value is on average 4.6% and 5.6% larger than the lower bound (not the optimal solution) obtained by the commercial solver for 80% and 95% decarbonization targets, respectively.

Acknowledgments

We thank John Parsons of MIT for our hydropower import data and for helpful discussion, suggestions and comments. We additionally thank Zack Schmitz of the MIT Energy Initative for his assistance with model runs and troubleshooting. We also thank NREL's ResStock team for assisting with the usage of the ResStock model. All authors acknowledge support from the MIT

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Parameter	Value
NG price ($/MMBtu)	5.45
LCF availability levels (MMBtu)	70.8e6, 214.5e6, 1000e6
LCF availability price ($/MMBtu)	10, 25, 50
Power load shedding cost ($/MWh)	10,000
Gas load shedding cost ($/MMBtu)	10,000
Reserve margin requirement	15%

Table 1: Key system parameter assumptions in JPoNG. The reserve margin is enforced for all modeled operational periods.

Energy Initiative Future Energy Systems Center. R.K. and S.A. acknowledge funding from MIT Climate Grand Challenges "Preparing for a new world of weather and climate extremes," and MIT Energy Initiative's project Hurricane Resilient Smart Grids. The responsibility for the contents lies with the authors.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Supplementary Information

6 Supplementary Results

7 Supplementary Results

7.1 Impacts of electrification on length of peak event

While peak electricity demand is one of the metrics commonly used in grid capacity planning, the length of the peak event is instrumental to determine the required mix of resources. Fig. S1 depicts the length of peak events during the summer and winter seasons for different electrification scenarios. We define a "peak event" as the continuous period for which the electricity demand is within 25% of the seasonal peak demand, including the periods directly before and after the peak hour.

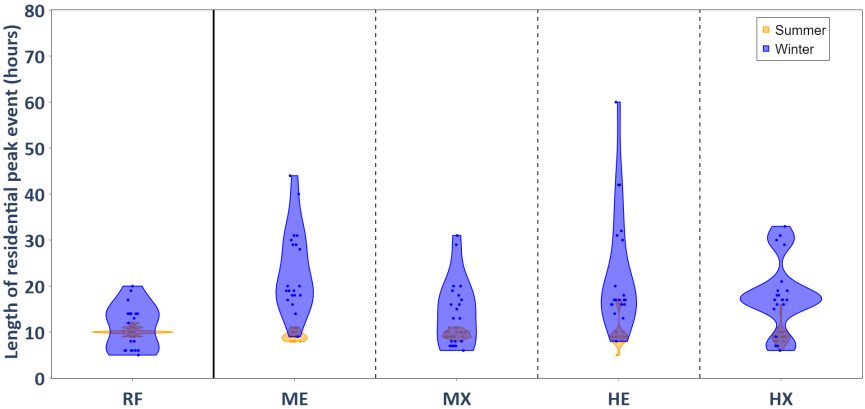


Fig. S1: Length of the residential peak electricity demand event for summer and winter across all 20 weather years for each electrification scenario. Length is defined as the number of continuous hours in the seasonal peak event for which the demand is greater than 75% of the peak seasonal demand. The width of the violins is proportional to the density of the data.

For certain weather years under certain electrification scenarios, the peak event can span multiple days, especially in the winter. When electrification is implemented in the absence of envelope improvements, peak events can become longer in duration. Envelope improvements can reduce the length of peak events.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

7.2 Demand implications of electric resistance backup

Our main analysis assumes that summer-sized systems will maintain their existing systems for backup heating. However, it is also possible that households will adopt electric resistance systems as their backup systems, for example, when their existing system eventually malfunctions and needs to be replaced. With a COP of nearly 1, electric resistance backup is inefficient in comparison to heat pumps. In Fig. S2, we highlight the peak demand impacts of using summer-sized heat pumps with electric resistance backup.

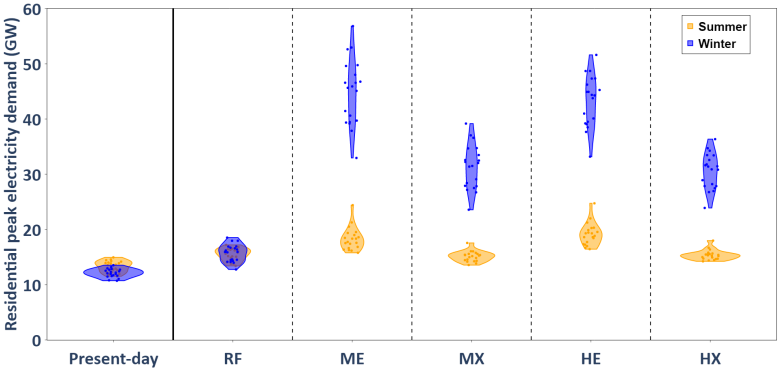


Fig. S2: Summer and winter peak loads under the electrification scenarios presented in the main text, but for which summer-sized systems use electric resistance backup rather than the existing system. Peak loads significantly increase.

Electric resistance instead of existing backup can further increase peak load between 0-184% for 2050 across scenarios, with large increases across the board for winter-peaking scenarios. Scenario ME, which emphasizes the deployment of summer-sized systems, now has the highest median peak of 45.7 GW, 25% higher than even the highest peak of 36.6 GW presented under the HE scenario in the main text. The sensitivity of peak load to interannual weather variation also increases drastically across the scenarios. For example, the range of peak demand under the ME scenario with electric resistance vs. existing backup is 33.0-56.8 GW vs. 16.3-20.0 GW, respectively. Therefore, the adoption of electric resistance may not be economical from a power system planning perspective.

7.3 Validation of model against historical data

Comparing our simulated present-day results to recent historical consumption values illustrates our model's performance. We include a comparison for the New England region in Fig. S3, which shows that our present-day model estimates the historical power demand similar to EIA data [58]. However, it generally overestimates residential gas demand, by 18-37% (average of 28%)

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

or 10.4 - 21.3 TWh (average of 16.7 TWh). Massachusetts contributes the majority of this demand error, by an average of 11.5 TWh, due to its large population and high proportion of homes with gas heating. The error for electricity demand is significantly smaller, with an error of 9-15% (average of 12%) across New England.

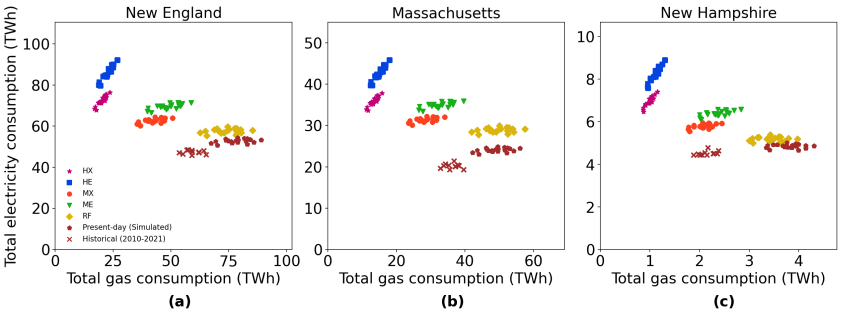


Fig. S3: Annual demand results including the historical EIA data and our present-day simulated results, indicated by the points in brown. All results shown use 400 archetypes per load zone. Results for Massachusetts and New Hampshire are highlighted to show how Massachusetts contributes an outsized portion of the error relative to other states. Note the difference in scale on the axes, preserved in order to enable visibility of the data clusters.

We identify three potential sources of error to explain the overestimation of gas demand relative to historical data from EIA [59] shown in Fig. S3:

1. Systematic errors intrinsic to the ResStock model.
2. Insufficient number of building archetypes in each zone.
3. Errors in weather data.

We also note that the EIA data itself may have errors relative to actual historical consumption – for example, some utilities report demand data to EIA in the form of billing periods rather than calendar months, which may misalign some of the demand data temporally [25]. However, because we are unable to quantify this error and others, we treat the EIA data as a satisfactory baseline for actual consumption.

Errors in ResStock model:

The calibration and validation of the ResStock model have been well-recorded by NREL [25], which used 550,000 archetypes nationally to analyze hourly building energy demands for a single weather year (2018). The subset of model outputs for the New England region compares reasonably well against EIA state-level monthly consumption data for gas and electricity (see Appendix F in the report [25]). For example, for the most populous state of Massachusetts, gas demand from NREL's ResStock results are 10% higher than the data from EIA (42.1 TWh vs. 38.2 TWh). Nonetheless, this deviation

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

between ResStock model outputs vs. EIA data also propogates into our own analysis using ResStock with fewer archetypes per state (see next section) and multiple weather years. In general, we opt not to attempt to calibrate ResStock to better fit historical demand. First, NREL already undertook a considerable effort to calibrate the model using immense amounts of proprietary and spatio-temporally resolved demand data [25], which we lack. Additionally, there would be no certainty that this would result in a more "accurate" result once the model is used to evaluate future levels of electrification that will fundamentally alter demand patterns from the historical baseline.

Number of building archetypes:

As observed in ResStock's performance, higher numbers of archetypes lead to smoothing and convergence in the demand outputs of ResStock [25]. Therefore, another source of the error can be attributed to the number of archetypes we consider in our bottom-up model, which in the main text is 400 archetypes per zone. Table S1 presents the impact of archetype count on Massachusetts, the most populous state in the region with the highest contribution to the gas demand overestimation.

Table S1: Impact of archetype count on annual present-day model error for 2018 in Massachusetts. Differences calculated relative to EIA empirical data and results from NREL's ResStock runs in the End Use Load Profiles Project [25]. Our demand analysis for all scenarios in the main text uses 400 archetypes per zone.

Archetypes per zone	Total archetypes	%diff. from NREL (gas)	%diff. from EIA (gas)	%diff. from NREL (power)	%diff. from EIA (power)
400	2,800	+17.5	+29.7	+0.92	+22.7
1,000	7,000	+12.1	+23.7	+3.46	+25.8
1,700	11,900	+10.3	+21.7	+4.26	+26.7
2,500	17,500	+7.5	+18.5	+4.84	+27.4

The results show that increasing the number of archetypes consistently reduces the gas demand error relative to EIA data, from an overestimate of 29.7% at the 400-archetypes-per-zone resolution used in our main analysis, down to 18.5% with 2,500 archetypes per zone. We expect the trend to continue with even higher archetype sizes as the quota sampling algorithm used in ResStock selects combinations of archetype features proportional to their probability distributions [25] and thus better captures outliers with higher archetype counts. We note that this observation is not aligned with NREL's report in which 1000 archetypes per zone of interest is recommended to obtain convergence in demand behavior [25]. However, the emphasis in that effort was calibrating electricity demand rather than gas demand; our findings indicate it may take greater numbers of samples to approach convergent gas demand results in ResStock.

As for the power demand, it can be seen that for Massachusetts, the model experiences a 22-28% deviation from EIA, which is outsized relative to the

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

rest of New England. However, this deviation is evidently due to intrinsic error in the ResStock model - our present-day model result for power demand never deviates further than 5% from NREL's estimate. We do not expect that we could improve the power demand results, given that this would require calibrating ResStock itself, as aforementioned.

Weather data:

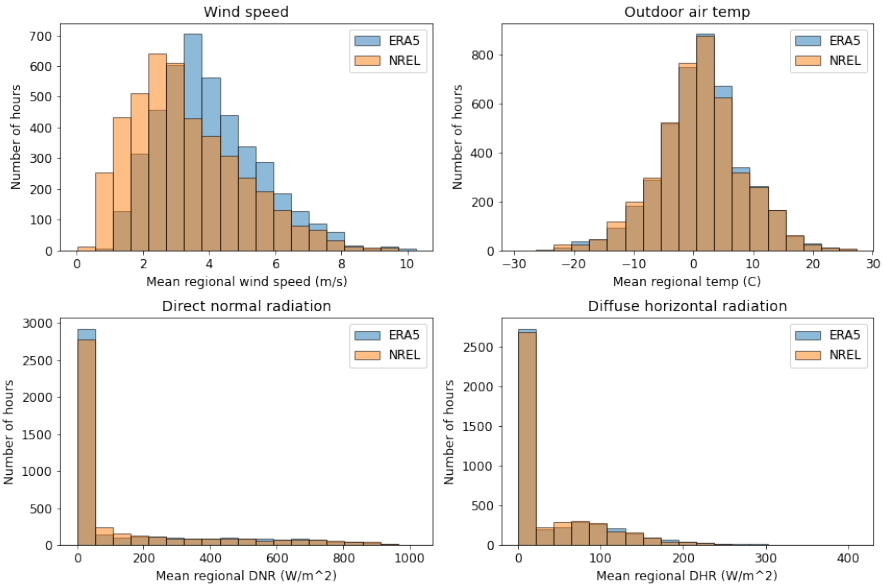


Fig. S4: Comparison between key weather features for the ERA5 and NREL weather datasets during the heating season (October - March). The hourly values are averaged across all of the 44 weather locations used in our models.

Because weather is also a key input to the bottom-up framework and significantly impacts the demand output, we sought to quantify its impact. NREL weather data encompasses the same parameters and geographic scope as our own weather data while producing gas demand outputs much closer to the EIA baseline; therefore, we compare our 2018 weather data to it in an attempt to determine what might be causing our overestimate. Furthermore, as mentioned in Section 12.4.1, the NREL weather largely is constructed from NOAA observational data for 2018, making it closer to a "ground truth" dataset. In particular, we compare key weather features in the heating season (defined here as October through March), which would be more likely to adversely affect gas demand because the majority of gas demand is due to space heating. We show comparison histograms for key weather features in Fig. S4. It should

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

be noted that the comparison of all weather-related parameters is not possible as NREL's published weather data purposefully excludes key variables like cloud cover and infrared radiation.

The figures reveal that while there is alignment between our ERA5 weather data and NREL's among most key variables, the wind speed has some discrepancies. When we weigh each location's weather data by its population to capture the likely impact on residential energy demand, we find that the ERA5 data generally overestimates winter wind speeds, by an average of 0.55 m/s (22%) annually. This value is perhaps substantial enough to appreciably change the demand outputs of the bottom-up model by increasing the demand for heating, as higher wind speeds increase building heat loss on cold days. We note that discrepancies with wind speeds in ERA5 have been noted by other researchers in this domain [60]. However, the complexity of the weather inputs coupled with the sophisticated impact of weather on heating demand in the bottom-up approach makes it difficult to concretely determine the key parameters that contribute most to the differences between the results of the two datasets.

7.4 Impact of historical gas demand overestimation on future demands

In the previous section, we identify a handful of potential input-related sources of the gas demand overestimation. Although increasing the archetype count can potentially reduce the error, it will greatly increase computational and data storage requirements, especially given the wide input parameter space of weather and electrification scenarios fundamental to this study. As explained in Section 12.4.1, we also opt to maintain our usage of the ERA5 data because it meets the needs of the bottom-up approach while performing well in our study region relative to other candidate datasets.

Importantly, the errors in the present-day analysis may not necessarily propagate to the future electrified scenarios that underlie the bulk of our findings. To see how the sources of gas demand overestimation impact future demand, we compare our results for New England with those of an alternatively parameterized model - an "NREL-parameterized" model - for a handful of our scenarios. The premise is to generate an instance that would reasonably approach the current "best" published output of the model by mirroring the setup NREL used in its extensive calibration efforts. The "NREL-parameterized" model setup has two differences from the model setup used to generate the results described in the main text. First, the number of archetypes in the "NREL-parameterized" case is the same as what NREL used in its calibration of ResStock for New England (26,701 total, divided among the states roughly proportional to the number of homes). Second, the weather data is the same as NREL's 2018 weather data, except we use TMY data to fill in the cloud cover data which was purposefully redacted, enabling ResStock to interpolate the remaining missing weather features. The rest of the model

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

inputs, including building stock evolution, population change, climate change, and electrification scenarios for 2050 are the same for both model setups.

Table S2: Percentage difference between electrification scenario results in the main analysis and "NREL-parameterized" results

Scenario	gas demand	power demand	power peak load
HX	+0.64	-4.95	-2.15
MX	+0.38	-5.79	-7.92
RF	+0.77	-7.52	-4.82

The results shown in Table S2 reveal that our future projections are more robust to the sources of overestimation impacting the present-day simulations. Across the HX, MX, and RF scenarios, annual gas demand is overestimated relative to the "NREL-parameterized" model setup by approximately just 0.38-0.77% as opposed to 17.5% for the present-day results. Annual electricity demand is also quite robust at only 4.95% to 7.52% lower than the "NREL-parameterized" results. Importantly, the difference in peak demand between the two parameterizations is small, with less than an 8% difference. The results indicate that using a larger archetype size or an alternative weather dataset would likely not confer a significant impact on our findings with respect to future supply-side investments. Furthermore, all the results in this discussion have pertained solely to the residential sector, which only makes up a portion of the economy-wide demand data fed through to our supply-side JPoNG model, therefore tempering the sensitivity of our model outputs to the residential demand alone.

7.5 Spatial Distribution of Peak Load Increases

In addition to the map in the main text, we include a tabular summary of the spatially diverse increases in peak load in Table S3 below.

7.6 Spatial Distribution of New Investment

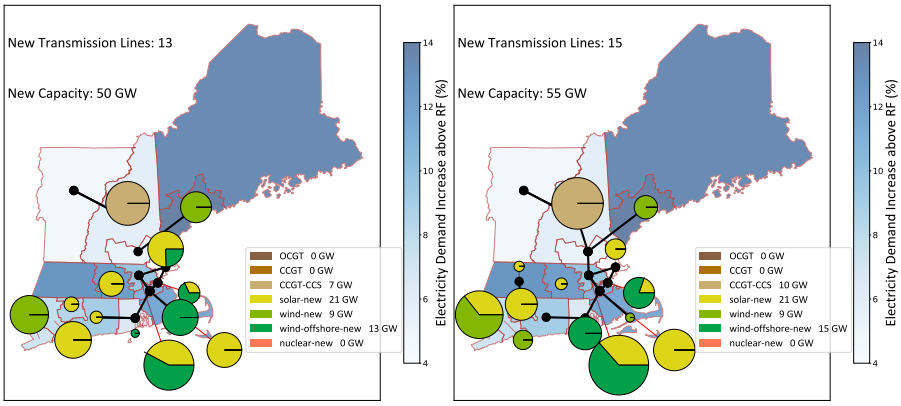
The spatial distribution of new generation and transmission capacity for the HX scenario under 80% and 95% decarbonization targets is highlighted in Fig. S5. As pointed out by previous studies [29, 61], the region needs substantial expansion of its current transmission line infrastructure to support a generation mix dominated by variable renewable energy (VRE) and achieve decarbonization targets. For the exemplar 2003 weather year, Fig. S5 shows significant investment for both decarbonization targets in new offshore wind and solar generation in coastal regions of Massachusetts.

The stringency of the decarbonization target necessitates investment in both new transmission lines and new power plants. As explained in Section 9.1.3, all candidate transmission lines are assumed to have the same maximum flow capacity. Across the two scenarios, the spatial distribution in

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Table S3: Average peak electricity demand increases by zone. Zone topology presented in Section 9.1.1.

Zone	Average Peak Electricity Demand Increase (%)	Average Peak Electricity Demand Increase (GW)
1	115.77	0.79
2	84.06	1.20
3	44.99	0.34
4	70.03	0.53
5	81.24	0.89
6	77.04	0.71
7	62.43	0.55
8	138.96	0.81
9	90.23	0.53
10	115.46	0.65
11	60.18	0.48
12	112.06	0.61
13	41.72	0.41
14	42.28	0.39
15	37.05	0.34
16	31.02	0.27
17	25.25	0.20



(a) 80% decarbonization target (b) 95% decarbonization target

Fig. S5: New transmission lines and capacity addition in the region under HX scenario and exemplar weather year of 2003. The size of the pie charts is proportional to total capacity it represents. The color intensity of each zone indicates the percentage of electricity load increase between HX and RF scenarios.

capacity does not strongly correlate with the spatial distribution in incremental load increases which can be due to transmission line expansion that connects the states with higher load increase, such as Maine, to states with higher capacity expansion, such as Massachusetts.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

7.7 Example Winter Day Operations

As an example of system operation under high electrification and decarbonization scenarios, Fig. S6 shows the various input and output indicators for four different winter days with the exemplar 2003 weather year and HX scenario under 80% and 95% decarbonization targets.

The total load generally follows the temperature; the total load is higher on colder "Day 3", and much lower on warmer "Day 1", "Day 2", and "Day 4." On the supply side, nuclear and domestic hydropower are operating close to nameplate capacity for all days, while CCGT-CCS is utilized for all four days at relatively high capacity factors except for "Day 1."In all four days, we see that battery storage and hydro imports from Canada (details in Section 9.1.2) play a critical role in meeting peak demand.

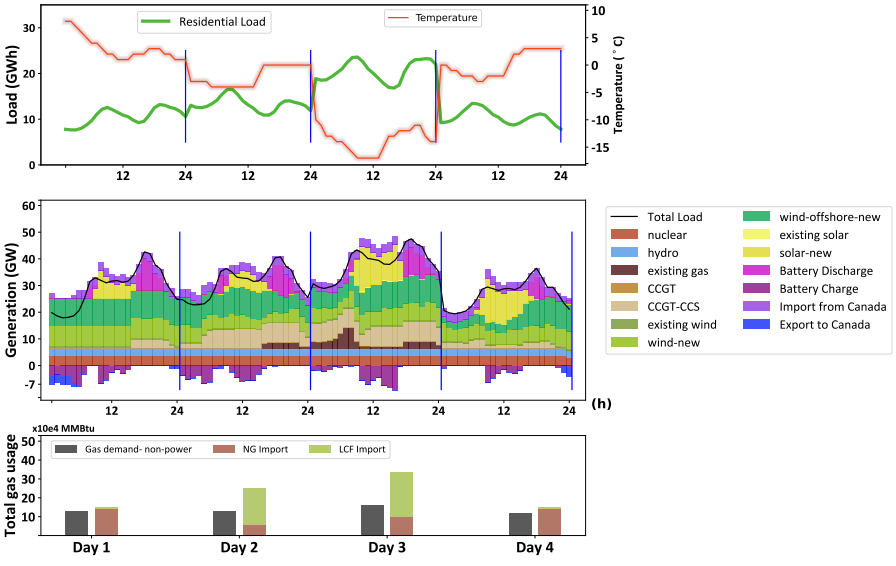
The generation mix depends on the total load and availability of VRE resources. "Day 1" has a low load, low solar CF but high wind CFs. Accordingly, the system fully deploys wind generation assets and compensates for peak load hours with battery storage. "Day 1" has negligible gas generation except for deployment of CCGT-CCS for several hours. Conversely, "Day 2" has significant generation from CCGT-CCS owing to relatively low solar CFs. The lower load of the day coupled with high wind CF, battery storage and import from Quebec is sufficient to meet the demand. "Day 3" is similar to "Day 2" in terms of wind CFs, but has a significantly higher load and solar CFs. A combination of high load due to the extreme cold temperatures and low wind CFs in the morning requires the dispatch of existing unabated gas generation. The load magnitude of "Day 4" is similar to that of "Day 2." However, the solar CFs are higher, but wind CFs are lower for the day, resulting in some deployment of abated gas generation under both decarbonization levels.

The imports of NG and LCF highly depend on gas generation plants in the power system. The total import and non-power demand for "Day 1" are almost equal under both decarbonization targets. The discrepancy between the non-power gas demand and gas import (i.e., NG and LCF) increases in other days (indicating gas is needed for power generation), culminating in "Day 3" where the total import is more than twice the non-power gas load. LCF is a more expensive gas fuel than NG, hence its import surges only when the model needs substantial gas generation but the decarbonization targets prevent the import of carbon-intensive NG beyond a certain level.

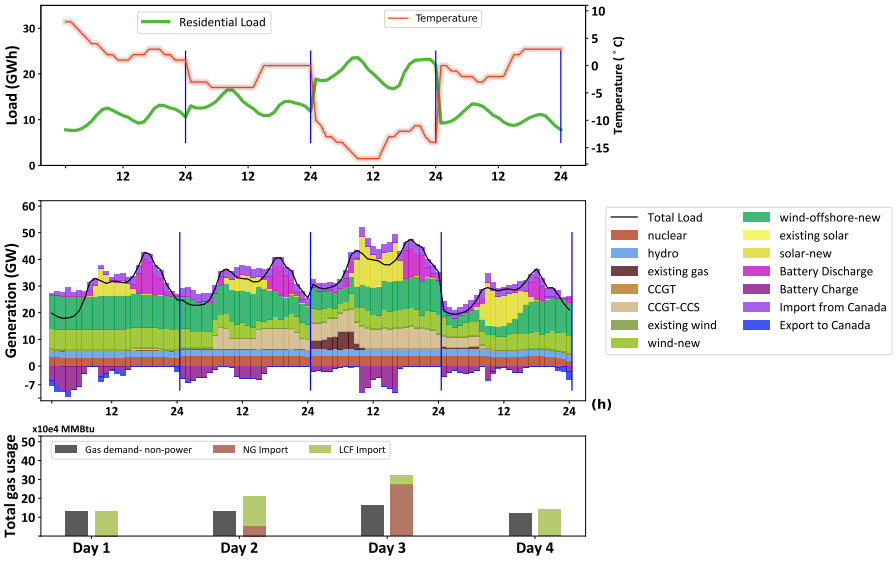
7.8 The Impact of CCS Technology

We explore the sensitivity of supply-side outcomes to the availability of CCS-based power generation, with results summarized Fig. S7, Fig. S8 and Table S4 Without CCS, the system cost increases by up to 4.2% in the 80% emissions case and by up to 5% in the 95% emissions reduction cases. This highlights the increasing value of low-carbon firm dispatchable generation sources like CCS-based power generation with increasing stringency of emissions constraint. Without CCS availability, the cost-optimal power capacity portfolio has a

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.



(a) 80% decarbonization target



(b) 95% decarbonization target

Fig. S6: Model inputs and outcomes for four representative winter days with exemplar 2003 weather year under HX scenario. The days are not necessarily sequential. The top plot shows the hourly residential load and ambient temperature on the left and right x-axes, respectively. The middle plot is the hourly dispatch decisions, battery charge and discharge, import from and export to Quebec, Canada, and the total hourly load of the region. The bottom plot depicts the daily gas load as well as daily NG and LCF imports.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

greater reliance on offshore wind, unabated gas power capacity which requires using more LCF for power generation (Table S4), and also battery energy storage capacity (primarily in the 95% emissions reduction case).

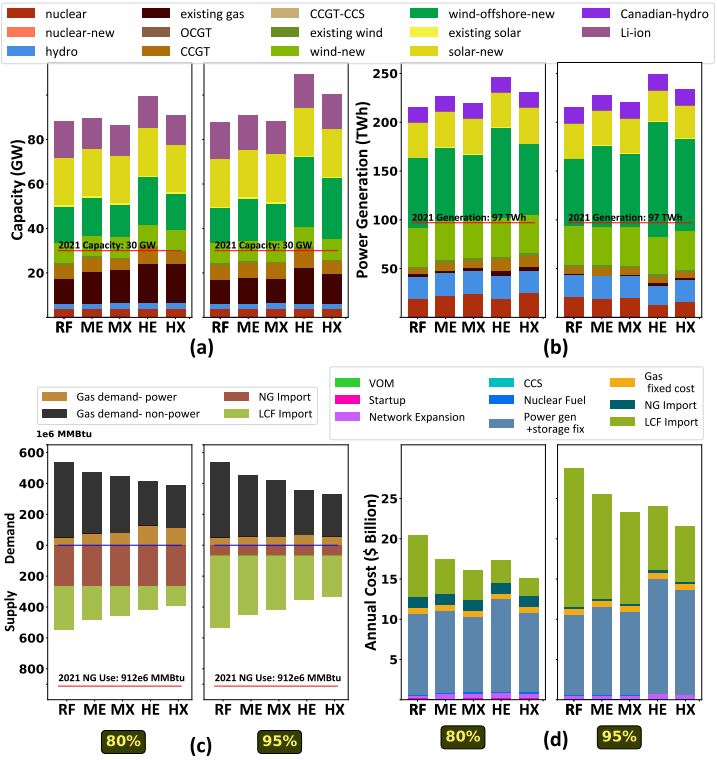


Fig. S7: Capacity, generation, gas consumption characteristics, and cost breakdown for the no-CCS case. In the 80% emissions cases, costs are comparable to the result where CCS is allowed. In the 95% case, costs are higher. Both emissions cases require higher amounts of offshore wind, unabated thermal, and LCF deployment than the corresponding case where CCS is allowed.

7.9 The impact of methane leakage

Our analysis in this paper is based on decarbonization goals which primarily target the emissions of carbon dioxide. However, methane emissions due to leakage in the supply chain of NG is currently substantial [62], though there are extensive retrofitting and mitigation programs that aim to reduce methane emissions from natural gas infrastructure in coming decades. For example, Massachusetts' Gas System Enhancement Program aims to repair leaks in the distribution infrastructure [63], while the EPA's new Methane Emissions Reduction program seeks to reduce upstream leakage [46].

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Table S4: Comparison of CCS and no-CCS cases for exemplar weather year 2003. Percent differences for each category span the five electrification scenarios presented in the main analysis. A positive percentage indicates a result that is higher for the no-CCS case.

Emissions reduction	Total cost	Offshore capacity	Unabated thermal cap.	LCF consumption	Storage capacity
80%	0-4.2 %	28-52%	27-64%	14-80%	3-12%
95%	2-5%	35-84%	32-64%	7-11%	3-14%

This section highlights the impact of methane emissions associated with supply chain of LCF and NG fuels on the supply-side outcomes, Specifically, we evaluate scenarios where we include methane emissions associated with natural gas supply chain (production, gathering, processing, transportation and storage) and LCF supply chain (transportation and storage).

While emissions from NG distribution and end-use are also prevalent, they remain a relatively small share of the overall leaks [62]. Moreover, as aforementioned, many regions including the state of Massachusetts have implemented pipe replacement programs which when completed would further reduce these leaks. Therefore, considering a 2050 scenario, we did not include distribution-related CH4 emissions. Recent estimates suggest that about 2.3% of gross gas production in the US is lost due to methane emissions from the production points to end-use [62], of which the upstream and midstream stages collectively are responsible for 96% of all emissions. For LCF, whose supply and upstream emissions are quite uncertain, we only account for midstream emissions, assuming that they will be similar to those for natural gas.

Table S5 summarizes the methane emissions intensity and associated kgCO2 equivalent emissions intensity for natural gas and LCF considered in the evaluated scenarios.

Fig. S9 shows the difference in outcomes for instances with and without methane leakage. For all electrification scenarios, incorporating methane emissions substitutes NG with lower emissions-intensive, but expensive, LCF, thus increasing the total system cost; For 80% and 95% decarbonization goals, the total system cost increases by up to 12% and 4%, respectively. The power sector generation impacts of accounting for methane emissions associated with NG and LCF in the emissions constraint is relatively minor as reflected by the y-axis scale of the generation and capacity charts.

7.10 The impact of transportation flexibility

This section evaluates the impact of flexibility on transportation load. Although our analysis is primarily focused on the impacts of heat electrification in the residential sector under varying future scenarios, transportation is another sector with substantial potential for increased load in coming decades. Furthermore, transportation load is oft-considered as the best candidate sector for demand flexibility interventions, enabled by various demand response programs and tariff designs.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

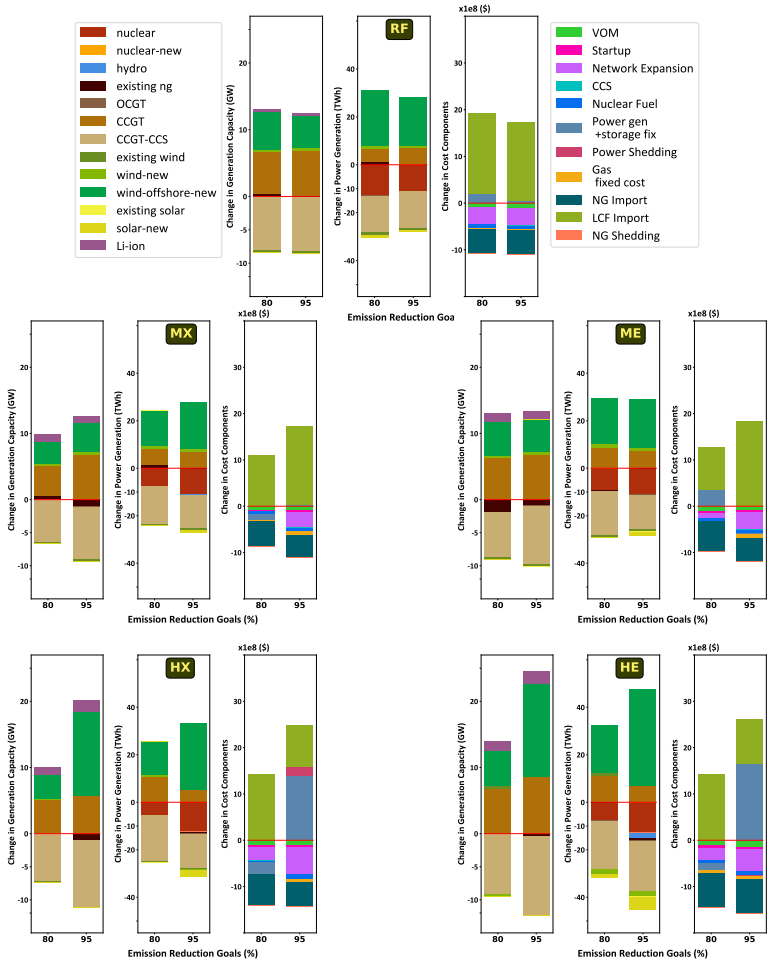


Fig. S8: Difference in capacity, generation, and cost components between with and without CCS instances for exemplar weather year 2003 under different decarbonization targets and electrification scenarios.

We used the modeling framework used in the GenX model [66] to formulate the representation of transportation flexibility. Let s_{nt}^{trRem} be the deferred demand remaining to be served during a flexibility event. Let $s_{nt}^{\text{ser}}, s_{nt}^{\text{def}}$ be served and deferred demand, respectively. We denote the number of hours the flexible transportation load can be advanced or delayed as T^{adv} and T^{def}. Let ρ_{nt}^{tr} be the share of the transportation load in the total load. The following constraints are added to the formulation:

$$s_{nt}^{\text{trRem}} = s_{n,t-1}^{\text{trRem}} + s_{nt}^{\text{def}} - s_{nt}^{\text{ser}} \qquad n \in \mathcal{N}^{\text{e}}, t \in \mathcal{T}^{\text{e}} \backslash \{t_{\tau}^{\text{start}} | \ \tau \in \mathfrak{R}\}$$

$$s_{nt}^{\text{def}} \leq \rho_{nt}^{\text{tr}} D_{n\phi_t^e}^{e} \qquad n \in \mathcal{N}^{\text{e}}, t \in \mathcal{T}^{\text{e}}$$

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

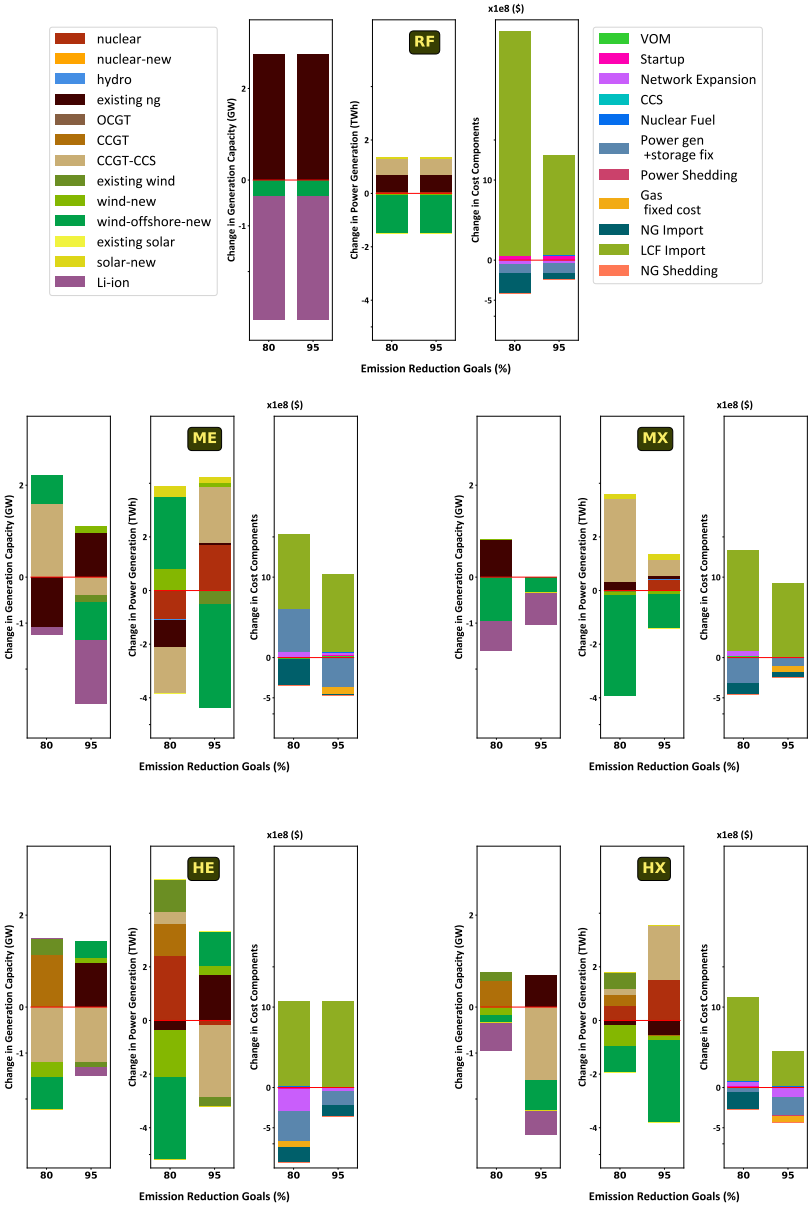


Fig. S9: Difference in capacity, generation, and cost components between with and without methane leakage instances (instances with methane leakage - without methane leakage) for exemplar weather year 2003 under different decarbonization targets and electrification scenarios.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Property	Value	Unit	Source
Volumetric energy density (HHV)	1089	btu/cf	[64]
Gravimetric density (32F and 1 atm)[1]	22	g/cf	[64]
Carbon content	72.40%	wt%	[64]
CO2 emissions intensity	0.0536	tCO2/MMBtu	
Pure methane properties			
Volumetric energy density (HHV)	1068	btu/cf	[64]
Gravimetric density (32F and 1 atm)[1]	20.3	g/cf	[64]
Carbon content	75.00%	wt%	[64]
CO2 emissions intensity	0.0523	tCO2/MMBtu	
Methane content	90%	vol%	
Methane leakage rate	2.30%	gCH4/gCH4 in NG	
Upstream emissions (production, gathering, processing, trans. + storage)[2]	96%		
Upstream emissions (transport + storage) of LCF[2]	14%		
100-year Global warming potential CH4	29.80	kgCO2eq/kgCH4	[65]
Upstream methane leakage rate of NG	0.022	gCH4/gCH4	
Midstream methane leakage rate of LCF	0.003	gCH4/gCH4	
Emissions leakage per unit of methane consumed	0.0226	gCH4/gCH4	
Total CO2eq emiss. intensity of NG	0.0649	tCO2eq/MMBtu	
Total CO2eq emiss. intensity of LCF	0.00161	tCO2eq/MMBtu	

Table S5: Methane emissions intensity for natural gas and LCF
[1] NG density at 60F and 1 atm [2] of total CH4 emissions [3] The capture rate for CCGT-CCS plants is reported for combustion and does not account for the upstream emissions. The inclusion of methane emissions increases the carbon intensity of NG by 22%. Therefore, we reduce the capture rate of these plants by 22%, to 69.3%.

$$\sum_{t'=t+1}^{t+T^{\mathrm{def}}} s^{\mathrm{ser}}_{nt'} \geq s^{\mathrm{trRem}}_{nt} \qquad n \in \mathcal{N}^{\mathrm{e}}, t \in \mathcal{T}^{\mathrm{e}}$$

$$\sum_{t'=t+1}^{t+T^{\mathrm{adv}}} s^{\mathrm{def}}_{nt'} \geq -s^{\mathrm{trRem}}_{nt} \qquad n \in \mathcal{N}^{\mathrm{e}}, t \in \mathcal{T}^{\mathrm{e}}$$

$$s^{\mathrm{trRem}}_{nt} \in \mathbb{R}, s^{\mathrm{def}}_{nt}, s^{\mathrm{ser}}_{nt} \in \mathbb{R}^{+} \qquad n \in \mathcal{N}^{\mathrm{e}}, t \in \mathcal{T}^{\mathrm{e}}$$

Consequently, we change the hourly power supply-demand constraints (7) to account for flexible demand from the transportation sector:

$$\sum_{i \in \mathcal{P}} p_{nti} + \sum_{m \in \mathcal{N}^{\mathrm{e}}} \sum_{l \in \mathcal{L}^{\mathrm{e}}_{nm}} \mathrm{sign}(n-m) f^{\mathrm{e}}_{\ell t} + \sum_{r \in \mathcal{S}^{\mathrm{e}}_{n}} (s^{\mathrm{eDis}}_{ntr} - s^{\mathrm{eCh}}_{ntr})$$
$$+ s^{\mathrm{def}}_{nt} - s^{\mathrm{ser}}_{nt} + a^{\mathrm{e}}_{nt} = D^{\mathrm{e}}_{n\phi^{\mathrm{e}}_t} + d_n E^{\mathrm{pipe}} \kappa^{\mathrm{pipe}}_n + E^{\mathrm{cprs}} E^{\mathrm{pump}} \kappa^{\mathrm{capt}}_{nt} \quad n \in \mathcal{N}^{\mathrm{e}}, t \in \mathcal{T}^{\mathrm{e}} \tag{2}$$

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

As indicated by other studies [31, 66], allowing the transportation load to be satisfied flexibly on an hour-to-hour basis allows for load-shifting within a day and thus reduces investment in Li-ion battery storage. We see a similar effect in our analysis, as shown in Fig. S10, which illustrates the difference in outcomes for instances with and without transportation flexibility for the case in which transportation load for light-duty vehicles can be shifted ahead or forward by five hours.
Across all residential sector demand and emissions scenarios, demand flexibility reduces investment in battery storage (by 40-60%) and power network expansion, and reduces overall gas consumption (see reduced biogas and NG import in cost charts). The reduction of battery storage as well as gas generation leads to total cost reduction of 3.1-6.3% across all instances. This reduction is more pronounced under more stringent decarbonization targets. Overall gas and LCF consumption decrease for all instances. However, the capacity of gas-fired plants increases in some scenarios to make up for the capacity contribution of battery storage resources to meet the capacity reserve margin constraint (Eq. 10a). As we did not consider the potential for demand flexibility to contribute to the resource adequacy (also called reserve margin), gas capacity has to make up for the reduced battery storage investment. Allowing demand flexibility to contribute to the capacity reserve margin constraint would presumably reduce gas capacity and further increase the cost savings from demand flexibility.

7.11 Supply-side outcomes for all 20 weather years

Fig. S11 shows the ranges for capacity of key power assets and consumption of gas for all 20 weather years for each emissions constraint. Each weather year is used to calculate both the residential demand and supply-side capacity factors.

Beyond the conclusions presented for the exemplary weather year of 2003 in the main text, a handful of high-level conclusions can be drawn from the figures. Optimal deployment of transmission, thermal, storage, and offshore wind capacity is sensitive to weather across both emissions cases. The amount of thermal capacity deployed increases at higher electrification levels. The 95% scenario requires much greater amounts of LCF than the 80% emissions case. However, the relative proportions of each resource deployed are generally consistent across weather years. Onshore wind and solar are deployed to the maximum possible extent in both emissions scenarios across all weather years due to the land availability constraint applied to the model.

7.12 NPV analysis of envelope improvements

In order to balance the bulk system savings of envelope improvements against their demand-side installation cost, we calculate an example from our model results. Our results indicate average bulk energy system-related savings of

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

roughly $2.04 billion annually in 2050 for HX over HE as a result of envelope improvements under the 95% decarbonization goal. About 55% of homes in the overall stock will receive envelope improvements in the HX scenario or about 3.87 million homes. If we assume the entry-level envelope improvements specified in this paper cost an average of $20,000 per home and that the costs of these envelope improvements are incurred at once in 2050, the net present cost in 2050 is approximately $77.4 billion. In order to assess the cost-benefit tradeoff in annual terms, we convert this capital cost into an annualized equivalent uniform series. We assume that the envelope improvements continue to deliver the same amount of bulk system savings for 40 years beyond 2050 and that the costs are distributed over the same time horizon. To account for a range of potential discount rates, we consider a 3.4% real discount rate in line with the DOE's methodology for determining the cost-effectiveness of building energy codes [67] and a 7.1% discount rate equivalent to that used in our supply-side modeling (see Table S11). The annualized cost is equal to the capital cost multiplied by the capital recovery factor as in Equation 3:

$$\$77.4bil \cdot CRF = \$77.4bil \cdot \frac{r(1+r)^{40}}{(1+r)^{40}-1} \tag{3}$$

For r=.034 and r=.071, the annualized cost is $3.57 billion and $5.87 billion respectively. The annual bulk system savings of the envelope improvements therefore offset approximately 34.8-57.1% of the envelope retrofit cost.

8 Joint Power-Gas Planning (JPoNG) Model Formulation

Our model, referred to as JPoNG, determines the minimum cost planning decisions for power and gas systems considering the two systems' interdependency. The proposed model considers a range of generation and storage technologies modeled via operational and policy constraints. The formulation allows different temporal resolutions for the operation of both systems since the data availability or planning requirements can be different for each system. For example, decisions related to power generation, such as dispatch amounts and unit commitment, require hourly resolution. However, gas system operation does not involve generation decisions and only deals with transmission and storage operations for which daily resolution may be sufficient. Moreover, due to the ability of gas pipelines to provide some storage via line packing, daily resolution for scheduling gas operations could facilitate the management of intra-day variations in gas demand. In the model, the operations of both systems are coupled through two sets of constraints. The first set ensures gas flow to the power system. The second coupling constraint limits the CO_2 emissions incurred by consuming fossil-derived NG in both power and gas systems.

We model power and gas system operation at an hourly and daily time resolution, respectively. For the power system, we only consider representative days of system operation to manage model tractability, while the operations

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

of the gas system are modeled across all days of the year. Moreover, we also model the import of electric power to the system by defining an import node in the power system. The details of these systems are provided in Section 9.1.2. The choice of daily resolution for the gas system is sufficient to capture macro-dynamics of gas flow and provides significant computational advantages [14]. Fig. S12 illustrates our modeling approach for the time resolutions of both systems.

The network representation in the model consists of three sets of nodes as depicted in Fig. S13. The first set represents power system nodes and is characterized by different generation technologies (i.e., plant types), demand, storage, and the set of adjacent nodes by which the node can exchange electricity. The only exception for power system nodes is the *import node* that, unlike other power nodes, operates on the entire year instead of representative days only. The details of the import node are provided in Section 9.1.2. The second set is gas nodes, each of which is 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 referred to as storage-vaporization-liquefaction (SVL) nodes. We allow for the possibility of gas storage infrastructure to be located far from demand or injection points in the network, as per existing practice [68] (see 9.2.1 for detailed discussion). Accordingly, our model makes a distinction between gas and SVL nodes to account for their distinct locations. We also allow the gas system to use LCF, which represents a renewable source of gas fuel that is interchangeable with NG and hence can be imported and transported by the gas pipelines [32]. For ease of exposition, we separately present the power and gas systems' models as well as coupling constraints. The full description of the mathematical notation used in the formulation is described in 8.1.

8.1 Nomenclature

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
$\mathcal{CCS} \subset \mathcal{P}$	gas-fired plant types with carbon capture technology
$\mathcal{H} \subset \mathcal{P}$	Thermal plant types
$\mathcal{Q} \subset \mathcal{P}$	Technology with a resource availability limit
$\mathcal{Q}'$	Set of technologies with resource availability limits
$\mathcal{T}^{\mathrm{e}}$	Index set of representative hours for power system
$\mathfrak{R}$	Representative days
$\mathfrak{T}^{\mathrm{e}}_{\tau}$	Hours in $\mathcal{T}^{\mathrm{e}}$ that are represented by day τ
$t^{\mathrm{start}}_{\tau}, t^{\mathrm{end}}_{\tau}$	First and last hour in $\mathfrak{T}^{\mathrm{e}}_{\tau}$

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

$\mathcal{L}^{\mathrm{e}}$	Existing and candidate transmission lines
$\mathcal{L}^{\mathrm{e}}_{nm}$	Existing and candidate transmission lines between node n and m
$\mathcal{S}^{\mathrm{e}}_{n}$	All energy storage systems types
$\mathcal{A}^{\mathrm{g}}_{n}$	Adjacent gas nodes for node n
$\mathcal{N}^{\mathrm{g}}, \mathcal{N}^{s}$	Gas and SVL nodes
$\mathcal{T}^{\mathrm{g}}$	Days of the planning year
$\mathcal{A}^{s}_{k}$	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
$\mathcal{L}^{\mathrm{LCF}}$	LCF availability levels and prices

Indices

n, m	Power system node
k	Gas system node
j	SVL facility node
i	Power generation plant type
r	Storage type for power network
ℓ	Electricity transmission line or gas pipeline
t	Time step for power system's operational periods
τ	Time step for gas system's operational periods
l	LCF availability or price level

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/hour]
$C^{\mathrm{pipeDec}}_{\ell}$	Decommissioning cost for pipeline ℓ [$/line]

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Annual Costs	
C_i^{fix}	Fixed operating and maintenance cost (FOM) for plants, [$]
C_r^{EnFix}	Energy-related FOM for storage, [$/MWh]
C_r^{pFix}	Power-related FOM for storage, [$/MW]
C_ℓ^{trFix}	Fixed cost of transmission line ℓ [$/line]
C_j^{strFix}	FOM for storage tanks, [$/MMBtu]
C_j^{vprFix}	FOM for vaporization plants, [$/MMBtu/hour]
C_ℓ^{pipeFix}	Fixed cost of pipeline ℓ [$/line]

Other Cost Parameters	
C_i^{var}	Variable operating and maintenance cost (VOM) for plants, [$/MWh]
C_i^{startUp}	Start-up cost for plants, [$]
C^{eShed}	Unsatisfied power demand cost, [$/MWh]
C_i^{fuel}	Fuel price for plants, [$/MMBtu]
C^{ng}	Fuel price for NG, [$/MMBtu]
C_l^{LCF}	Price of LCF at availability interval $[U_l^{\text{LCF}}, U_{l-1}^{\text{LCF}}]$ [$/MMBtu]
C^{gShed}	Unsatisfied gas demand cost [$/MMBtu]

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Other Parameters for the Power System

ρ_{nti}	Capacity (availability) factor for renewable plants
D^{e}_{nt}	Power demand, [MWh]
h_i	Heat rate, [MMBtu/MWh]
b_ℓ	Susceptance of line $\ell \in \mathcal{L}^{\mathrm{e}}$
η_i	Carbon capture rate, [%]
U^{prod}_i	Nameplate capacity, [MW]
L^{prod}_i	Minimum stable output, [%]
U^{ramp}_i	Ramping limit, [%]
γ^{eCh}_r	Charge rate for storage
γ^{eDis}_r	Discharge rate for storage
γ^{loss}_r	hourly self-discharge rate for storage
I^{trans}_ℓ	Initial capacity for transmission line ℓ, [MW]
U^{trans}_ℓ	Upper bound for capacity of transmission line ℓ, [MW]
$\mathcal{I}^{\mathrm{trans}}_\ell$	1, if trans. line ℓ exists; 0, otherwise
I^{num}_{ni}	Initial number of plants
$U^{\mathrm{e}}_{\mathrm{emis}}$	Baseline emission of CO_2 in 1990 from generation consumption, [ton]
U^{CCS}	Total annual carbon storage capacity, [ton]
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]
$U^{\mathrm{prod}}_{\mathcal{Q}}$	Production capacity for set of plants $\mathcal{Q} \subset \mathcal{P}$, [MW]
ζ	emissions reduction goal
w_t	Weight of the representative period t
ϕ^{e}_t	Mapping of representative period t to its original period in the time series
R^{CRM}	Capacity reserve margin rate
$\gamma^{\mathrm{CRM}}_{nit}$	Capacity derating factor of plant type i at node n at time t

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Import Node Parameters	
γ_t^{inflow}	inflow rate to reservoir [MWh/h]
U^{eCap}	Maximum energy capacity of reservoir [MWh]
U^{powCap}	Maximum power capacity of reservoir [MWh]
C^{fixImp}	FOM cost of power capacity [$/MWh/year]

Other Parameters for the Gas System	
$D_{k\tau}^{\text{g}}$	Gas demand, [MMBtu]
η^{g}	Emission factor for NG [ton CO_2/MMBtu]
U_k^{inj}	Upper bound for gas supply, [MMBtu]
γ_j^{liqCh}	Charge efficiency of liquefaction plant
γ_j^{vprDis}	Discharge efficiency of vaporization plant
β	Boil-off gas coefficient
I_ℓ^{pipe}	Initial capacity for pipeline ℓ, [MMBtu/day]
U_ℓ^{pipe}	Upper bound capacity for pipeline ℓ, [MMBtu/day]
$\mathcal{I}_\ell^{\text{pipe}}$	1, if the pipeline ℓ exists; 0, otherwise
I_j^{gStr}	Initial storage capacity, [MMBtu]
I_j^{vpr}	Initial vaporization capacity, [MMBtu/d]
I_j^{liq}	Initial liquefaction capacity, [MMBtu/d]
I_{kj}^{store}	Initial capacity of storage facility
U_l^{LCF}	Available levels for LCF, [MMBtu]
$U_{\text{emis}}^{\text{g}}$	Baseline emission of CO_2 in 1990 from non-generation consumption, [ton]
Ω_n	representative day for day n

Investment Decision Variables	
$x_{ni}^{\text{op}} \in \mathbb{Z}^+$	Number of available thermal plants $i \in \mathcal{H}$
$x_{ni}^{\text{op}} \in \mathbb{R}^+$	Number of available VRE plants $i \in \mathcal{R}$
$x_{ni}^{\text{est}} \in \mathbb{Z}^+$	Number of new thermal plants established, $i \in \mathcal{H}$
$x_{ni}^{\text{est}} \in \mathbb{R}^+$	Number of new VRE plants established, $i \in \mathcal{R}$
$x_{ni}^{\text{dec}} \in \mathbb{Z}^+$	Number decommissioned thermal plants, $i \in \mathcal{H}$
$x_{ni}^{\text{dec}} \in \mathbb{R}^+$	Number decommissioned VRE plants, $i \in \mathcal{R}$
$y_{nr}^{\text{eCD}} \in \mathbb{R}^+$	Charge/discharge capacity of storage battery
$y_{nr}^{\text{eLev}} \in \mathbb{R}^+$	Battery storage level
$z_\ell^{\text{eInv}} \in \mathbb{B}$	1, if transmission line ℓ is built; 0, otherwise
$z_\ell^{\text{gInv}} \in \mathbb{B}$	1, if pipeline ℓ is built; 0, otherwise
$z_\ell^{\text{gDec}} \in \mathbb{B}$	1, if pipeline ℓ is decommissioned; 0, otherwise
$z_\ell^{\text{gOp}} \in \mathbb{B}$	1, if pipeline ℓ is operational; 0, otherwise

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Other Decision Variables for Power System	
$p_{nti} \in \mathbb{R}^+$	Generation rate, [MW]
$x_{nti} \in \mathbb{R}^+$	Number of committed plants
$x^{\text{down}}_{nti} \in \mathbb{R}^+$	Number of plants shut-down
$x^{\text{up}}_{nti} \in \mathbb{R}^+$	Number of plants started up
$f^{\text{e}}_{\ell t} \in \mathbb{R}$	Flow rates, [MW]
$\theta_{nt} \in \mathbb{R}$	Phase angle
$s^{\text{eCh}}_{ntr} \in \mathbb{R}^+$	Storage charged, [MW]
$s^{\text{eDis}}_{ntr} \in \mathbb{R}^+$	Storage discharged, [MW]
$s^{\text{eLev}}_{ntr} \in \mathbb{R}^+$	Storage level, [MWh]
$s^{\text{rem}}_{n\tau r} \in \mathbb{R}$	Storage carry over during day τ for storage type $r \in \mathcal{S}^{\text{sL}}$
$s^{\text{day}}_{n\tau r} \in \mathbb{R}^+$	Storage level at the beginning of day τ for storage type $r \in \mathcal{S}^{\text{sL}}$
$\kappa^{\text{capt}}_{nt} \in \mathbb{R}^+$	Captured CO_2 [ton/h]
$\kappa^{\text{pipe}}_{n} \in \mathbb{R}^+$	CO_2 pipeline capacity [ton/h]
$a^{\text{e}}_{nt} \in \mathbb{R}^+$	Amount of load shedding, [MWh]
$\mathcal{E}^{\text{e}}$	Total emission from power system

Import Node Variables	
$s^{\text{cap-imp}}_{t}$	Energy capacity of reservoir
x^{imp}	Maximum power capacity

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Other Decision Variables for NG System (all in MMBtu)	
$x_j^{\mathrm{gStr}} \in \mathbb{R}^+$	Installed additional storage capacities
$x_j^{\mathrm{vpr}} \in \mathbb{R}^+$	Installed additional vaporization capacities
$f_{\ell\tau}^{\mathrm{g}} \in \mathbb{R}^+$	Flow rates
$f_{kn\tau}^{\mathrm{ge}} \in \mathbb{R}^+$	Flow rates from gas nodes to power nodes
$f_{kj\tau}^{\mathrm{gl}} \in \mathbb{R}^+$	Flow rates from gas nodes to liquefaction plants
$f_{jk\tau}^{\mathrm{vg}} \in \mathbb{R}^+$	Flow rates from vaporization plants to gas nodes
$g_{k\tau} \in \mathbb{R}^+$	gas supply (injection)
$s_{j\tau}^{\mathrm{gStr}} \in \mathbb{R}^+$	Storage capacities
$s_{j\tau}^{\mathrm{vpr}}, s_{j\tau}^{\mathrm{liq}} \in \mathbb{R}^+$	Vaporization and liquefaction amounts
$a_{k\tau}^{\mathrm{g}} \in \mathbb{R}^+$	Amount of load shedding
$a_{k\tau l}^{\mathrm{LCF}} \in \mathbb{R}^+$	Amount of LCF consumed
$y_l^{\mathrm{LCF}} \in \mathcal{B}$	1, if LCF consumption exceeds level l
$\lambda_l^{\mathrm{LCF}} \in [0,1]$	Share of LCF consumption at interval $[U_l^{\mathrm{LCF}}, U_{l-1}^{\mathrm{LCF}}]$
$\mathcal{E}^{\mathrm{g}}$	Total emission from gas system

8.2 Power System Model

Objective Function:

$$\begin{aligned}
\min \sum_{n\in\mathcal{N}^{\mathrm{e}}}\sum_{i\in\mathcal{P}}(C_i^{\mathrm{inv}}x_{ni}^{\mathrm{est}} + C_i^{\mathrm{fix}}x_{ni}^{\mathrm{op}} + \sum_{r\in\mathcal{S}_n^{\mathrm{e}}}(C_r^{\mathrm{pInv}} + C^{\mathrm{pFix}})y_{nr}^{\mathrm{eCD}}) \\
+ \sum_{n\in\mathcal{N}^{\mathrm{e}}}\sum_{r\in\mathcal{S}_n^{\mathrm{e}}}(C_r^{\mathrm{EnInv}} + C^{\mathrm{EnFix}})y_{nr}^{\mathrm{eLev}} && \text{(4a)} \\
+ \sum_{n\mathcal{N}^{\mathrm{e}}}\sum_{i\in\mathcal{P}} C_i^{\mathrm{dec}}x_{ni}^{\mathrm{dec}} && \text{(4b)} \\
+ \sum_{n\in\mathcal{N}^{\mathrm{e}}}\sum_{i\in\mathcal{P}}\sum_{t\in\mathcal{T}^{\mathrm{e}}} w_t p_{nti} C_i^{\mathrm{var}} && \text{(4c)} \\
+ \sum_{n\in\mathcal{N}^{\mathrm{e}}}\sum_{t\in\mathcal{T}^{\mathrm{e}}}\sum_{i\in\mathcal{P}} w_t C_i^{\mathrm{startUp}} x_{nti}^{\mathrm{up}} && \text{(4d)} \\
+ \sum_{l\in\mathcal{L}^{\mathrm{e}}} C_\ell^{\mathrm{trans}} z_\ell^{\mathrm{eInv}} \\
+ \sum_{\ell\in\mathcal{L}^{\mathrm{e}}:I_\ell^{\mathrm{trans}}=1} C_\ell^{\mathrm{trFix}} + \sum_{\ell\in\mathcal{L}^{\mathrm{e}}:I_\ell^{\mathrm{trans}}=0} C_\ell^{\mathrm{trFix}} z_\ell^{\mathrm{eInv}} && \text{(4e)} \\
+ \sum_{n\in\mathcal{N}^{\mathrm{e}}} d_n C_{\mathrm{CO_2}}^{\mathrm{inv}} \kappa_n^{\mathrm{pipe}} + C_{\mathrm{CO_2}}^{\mathrm{str}} \sum_{n\in\mathcal{N}^{\mathrm{e}}}\sum_{t\in\mathcal{T}^{\mathrm{e}}} w_t \kappa_{nt}^{\mathrm{capt}} && \text{(4f)} \\
+ \sum_{n\in\mathcal{N}^{\mathrm{e}}}\sum_{i\in\mathcal{P}}\sum_{t\in\mathcal{T}^{\mathrm{e}}} w_t p_{nti}(C_i^{\mathrm{fuel}} h_i) && \text{(4g)}
\end{aligned}$$

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

$$+ \sum_{n \in \mathcal{N}^{\mathrm{e}}} \sum_{t \in \mathcal{T}^{\mathrm{e}}} w_t C_n^{\mathrm{eShed}} a_{nt}^{\mathrm{e}} \tag{4h}$$

$$+ C^{\mathrm{fixImp}} x^{\mathrm{imp}} \tag{4i}$$

The objective function (4) minimizes the total investment and operating costs incurred in the power system. The first term (4a) is the investment and fixed operation and maintenance (FOM) costs for generation and storage. The term (4b) captures the cost of plant retirement or decommissioning. The variable operating and maintenance (VOM) cost is represented by the term (4c). The fourth term (4d) corresponds to startup costs of power plants with unit commitment constraints (i.e. thermal power plants). The network expansion and FOM costs are included in term (4e). The cost of CO_2 transport and storage infrastructure required to accompany CCGT-CCS power generation is incorporated by term (4f) which also captures the cost associated with establishing CO_2 pipelines and storage. Here, 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 CO_2 transport. The cost of gas consumption for non NG-fired power plants (i.e., nuclear) is ensured by term (4g). The term (4h) penalizes the load shedding in the power system. The last term (4i) incurs the fixed cost for the power capacity of the import node. As mentioned, we only report the costs associated with non-import power node.

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

$$x_{ni}^{\mathrm{op}} = I_{ni}^{\mathrm{num}} - x_{ni}^{\mathrm{dec}} + x_{ni}^{\mathrm{est}} \tag{5a}$$

$$x_{nti} - x_{n,t-1,i} = x_{nti}^{\mathrm{up}} - x_{nti}^{\mathrm{down}} \qquad t \in \mathcal{T}^{\mathrm{e}} \tag{5b}$$

$$x_{nti} \leq x_{ni}^{\mathrm{op}} \qquad t \in \mathcal{T}^{\mathrm{e}} \tag{5c}$$

Constraints (5a) model the number of operating plants. The unit commitment constraints for each node and plant type are presented in (5b), which computes the number of plants committed, started up, or shut down across the hours of the representative days. Constraints (5c) limit the number of committed units to the number of available ones. As per other similar studies [15, 57, 69, 70], we relaxed the integrality of unit commitment decisions which has been shown to introduce relatively small numerical error while significantly reducing the computational complexity of the problem.

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

$$L_i^{\mathrm{prod}} x_{nti} \leq p_{nti} \leq U_i^{\mathrm{prod}} x_{nti} \qquad i \in \mathcal{H} \tag{6a}$$

$$|p_{nti} - p_{n,(t-1),i}| \leq U_i^{\mathrm{ramp}} U_i^{\mathrm{prod}} (x_{nti} - x_{nti}^{\mathrm{up}}) + \max(L_i^{\mathrm{prod}}, U_i^{\mathrm{ramp}}) U_i^{\mathrm{prod}} x_{nti}^{\mathrm{up}} \qquad i \in \mathcal{H} \tag{6b}$$

$$p_{nti} \leq \rho_{nti} U_i^{\mathrm{prod}} x_{ni}^{\mathrm{op}} \qquad i \in \mathcal{R} \tag{6c}$$

$$a_{nt}^{\mathrm{e}} \leq D_{n\phi_t^{\mathrm{e}}}^{\mathrm{e}} \tag{6d}$$

The generation limits are imposed in constraints (6a). Constraints (6b) 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 (6c) limit the generation of VRE to hourly capacity factor (i.e. ρ_{nti}) of maximum available capacity (i.e. $U_i^{\mathrm{prod}} x_{ni}^{\mathrm{op}}$). Constraints (6d) state that the load-shedding amount can not exceed demand. Note that we use the mapping ϕ_t^{e} to access the demand in the corresponding hour of a representative day.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

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

$$\sum_{i\in\mathcal{P}} p_{nti} + \sum_{m\in\mathcal{N}^{\mathrm{e}}} \sum_{l\in\mathcal{L}^{\mathrm{e}}_{nm}} \mathrm{sign}(n-m) f^{\mathrm{e}}_{\ell t} + \sum_{r\in\mathcal{S}^{\mathrm{e}}_n} (s^{\mathrm{eDis}}_{ntr} - s^{\mathrm{eCh}}_{ntr}) + a^{\mathrm{e}}_{nt} = D^{\mathrm{e}}_{n\phi^{\mathrm{e}}_t} + d_n E^{\mathrm{pipe}} \kappa^{\mathrm{pipe}}_n + E^{\mathrm{cprs}} E^{\mathrm{pump}} \kappa^{\mathrm{capt}}_{nt} \tag{7}$$

Constraints (7) ensure that for each node and for each planning period the generation, the net flow, the net storage, and the load shedding amount should be 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 $\mathrm{sign}(n-m)$ 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 $f^{e}_{\ell t}$ appears with opposite signs (i.e., negative of positive signs) in the balance equations of the nodes connected by transmission line ℓ.

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

$$|f^{\mathrm{e}}_{lt}| \le I^{\mathrm{trans}}_{\ell} \qquad \text{if } \mathcal{I}^{\mathrm{trans}}_{\ell} = 1 \tag{8a}$$

$$|f^{\mathrm{e}}_{lt}| \le U^{\mathrm{trans}}_{\ell} z^{\mathrm{eInv}}_{\ell} \qquad \text{if } \mathcal{I}^{\mathrm{trans}}_{\ell} = 0 \tag{8b}$$

$$|f^{\mathrm{e}}_{\ell t} - b_{\ell}(\theta_{mt} - \theta_{nt})| \le M(1 - z^{\mathrm{eInv}}_{\ell}) \qquad \text{if } \mathcal{I}^{\mathrm{trans}}_{\ell} = 0 \tag{8c}$$

$$f^{\mathrm{e}}_{\ell t} = b_{\ell}(\theta_{mt} - \theta_{nt}) \qquad \text{if } \mathcal{I}^{\mathrm{trans}}_{\ell} = 1 \tag{8d}$$

$$\theta_{0,t} = 0 \tag{8e}$$

Flow for the existing transmission lines is limited by constraints (8a). Constraints (8b) limit the flow in candidate transmission lines only if it is already established (i.e., z^{t}_{ℓ}=1). Throughout the paper, we use M to denote a big number. Direct Current (DC) power flow constraints for the candidate and existing transmission lines are respectively imposed in constraints (8c) and (8d). The phase angle for the reference node 0 is set in constraint (8e).

Storage Constraints: For every $n \in \mathcal{N}^{\mathrm{e}}, r \in \mathcal{S}^{\mathrm{e}}_n$

$$s^{\mathrm{eLev}}_{nt^{\mathrm{start}}_{\tau}r} = (1-\gamma^{\mathrm{loss}}_r)(s^{\mathrm{eLev}}_{n,t^{\mathrm{end}}_{\tau}r} - s^{\mathrm{rem}}_{n\tau r}) + \gamma^{\mathrm{eCh}}_r s^{\mathrm{eCh}}_{nt^{\mathrm{start}}_{\tau}r} - \frac{s^{\mathrm{eDis}}_{nt^{\mathrm{start}}_{\tau}r}}{\gamma^{\mathrm{eDis}}_r}, \qquad \tau \in \mathfrak{R} \tag{9a}$$

$$s^{\mathrm{eLev}}_{n,t-1,r} = (1-\gamma^{\mathrm{loss}}_r)(s^{\mathrm{eLev}}_{ntr}) + \gamma^{\mathrm{eCh}}_r s^{\mathrm{eCh}}_{ntr} - \frac{s^{\mathrm{eDis}}_{ntr}}{\gamma^{\mathrm{eDis}}_r} \qquad t \in \mathcal{T}^{\mathrm{e}} \backslash \{t^{\mathrm{start}}_{\tau} |\ \tau \in \mathfrak{R}\} \tag{9b}$$

$$s^{\mathrm{day}}_{n,\tau+1,r} = (1-24\gamma^{\mathrm{loss}}_r) s^{\mathrm{day}}_{n,\tau,r} + s^{\mathrm{rem}}_{n\Omega_{\tau},r}, \quad \tau \in \mathcal{T}^{\mathrm{g}} \backslash 365 \tag{9c}$$

$$s^{\mathrm{day}}_{n,1,r} = (1-24\gamma^{\mathrm{loss}}_r) s^{\mathrm{day}}_{n,\tau,r} + s^{\mathrm{rem}}_{n\Omega_{\tau},r}, \quad \tau = 365 \tag{9d}$$

$$s^{\mathrm{day}}_{n\tau r} = s^{\mathrm{eLev}}_{nt^{\mathrm{end}}_{\tau}r} - s^{\mathrm{rem}}_{n\tau r}, \quad \tau \in \mathfrak{R} \tag{9e}$$

$$s^{\mathrm{rem}}_{n\tau r} = 0, \quad \tau \in \mathfrak{R}, r \in \mathcal{S}^{\mathrm{sS}} \tag{9f}$$

$$s^{\mathrm{eCh}}_{ntr} \le y^{\mathrm{eCD}}_{nr} \tag{9g}$$

$$s^{\mathrm{eCh}}_{ntr} \le y^{\mathrm{eCD}}_{nr} \tag{9h}$$

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

$$s^{\text{eLev}}_{ntr} \leq y^{\text{eLev}}_{nr} \tag{9i}$$

Recall (see Fig. S12) that representative days are not necessarily consecutive; hence our formulation accounts for the carryover storage level between representative days, which is particularly important when modeling LDES. Li et al. [70] enforce the beginning and ending storage levels of each representative day to 50% of the maximum storage level. Here, we use a similar approach for short-duration batteries in which we time-wrap the beginning and ending hours of a day. That is, we assume the same charging state for the beginning and ending hours of a day, implicitly precluding energy carryover between representative days. For LDES, however, we use the method proposed in [66, 71] in which the unrestricted variable $s^{\text{rem}}_{n\tau r}$ models the carryover from a representative day τ to the next.

Constraints (9a) model battery storage dynamics for the initial hours of each representative day. Constraints (9b) model the storage balance for the remaining hours. The energy transfer across two consecutive representative days is modeled via constraints (9c). The storage in the first and last representative days is related by constraints (9d). Constraint (9e) only applies to representative days and ensures that the storage at the beginning of a day is equal to the storage level at the last hour of the day minus the storage carryover. The storage carryover for short-duration batteries is prevented by constraints (9f). Finally, the charge/discharge limits on storage level are imposed in constraints (9g) to (9i). Analogous to similar studies on power system expansion [57, 70], we do not account for use-dependent storage capacity degradation.

Resource Availability Constraints:

$$\sum_{n\in\mathcal{N}^{\text{e}}}\sum_{i\in\mathcal{Q}} U^{\text{prod}}_{i} x^{\text{op}}_{ni} \leq U^{\text{prod}}_{\mathcal{Q}} \qquad \mathcal{Q}\in\mathcal{Q}' \tag{10}$$

We consider resource availability limits for the development of VRE sources. In comparison to thermal plants, the siting of renewable resources is a major challenge due to the relatively large land area footprint per MW, the spatial heterogeneity in their resource availability and land availability limits due to such non-energy considerations as preserving the natural landscape [29]. Therefore, constraint (10) limits the installed capacity of a certain set of power plants to their maximum availability limit. The parameter $\mathcal{Q}$ denotes a generation technology class for which there is a resource availability limit. These classes include solar, onshore wind, offshore wind, and nuclear and are represented by set $\mathcal{Q}'$. Note that each technology class can include multiple plant types. For example, nuclear technology can include existing and new nuclear plant types.

Carbon Capture and Storage (CCS) Constraints: For every $n\in\mathcal{N}^{\text{e}}, t\in\mathcal{T}^{\text{e}}$

$$\kappa^{\text{capt}}_{nt} = \eta^{\text{g}}\eta_i h_i p_{nti} \qquad i\in\mathcal{CCS} \tag{11a}$$

$$\kappa^{\text{capt}}_{nt} \leq \kappa^{\text{pipe}}_{n} \tag{11b}$$

$$\sum_{n\in\mathcal{N}^{\text{e}}}\sum_{t\in\mathcal{T}^{\text{e}}} \kappa^{\text{capt}}_{nt} \leq U^{\text{CCS}} \tag{11c}$$

The constraint (11a) computes the amount of captured carbon in gas-fired power plants equipped with CCS technology. Constraint (11b) determines the CO_2 pipeline capacity. Finally, constraint (11c) limits the total amount of captured CO_2 to the annual CO_2 storage capacity.

Capacity Reserve Margin (CRM):

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

$$\sum_{n\in\mathcal{N}^{\mathrm{e}}}\left(\sum_{i\in\mathcal{P}}\gamma_{nit}^{\mathrm{CRM}}U_i^{\mathrm{prod}}x_{ni}^{\mathrm{op}}+\sum_{r\in\mathcal{S}_n^{\mathrm{e}}}(s_{ntr}^{\mathrm{eDis}}-s_{ntr}^{\mathrm{eCh}})\right)\geq(1+R^{\mathrm{CRM}})\sum_{n\in\mathcal{N}^{\mathrm{e}}}D_{n\phi_t^{\mathrm{e}}}^{\mathrm{e}},\quad t\in\mathcal{T}^{\mathrm{e}} \tag{12a}$$

The CRM constraint ensures that the installed capacity in the system plus the net discharging power from storage technologies exceeds a certain level of aggregated load across the region for all time periods. CRM is usually defined for the peak hour load, but we impose it for all time periods as our model is intended to be used for future infrastructure planning where the load projections are uncertain even if they are modeled as deterministic parameters.

Import node constraints: Assuming the n' is the import node, the set of associated constraints are:

$$s_0^{\text{cap-imp}}=0.7U^{\mathrm{eCap}} \tag{13a}$$

$$s_t^{\text{cap-imp}}=s_{t-1}^{\text{cap-imp}}+\gamma^{\text{inflow}}-p_{n'ti}\qquad \mathfrak{T}_\tau^e,\tau\in\mathcal{T}^g,i\in\mathcal{P} \tag{13b}$$

$$s_{8760}^{\text{cap-imp}}=0.7U^{\mathrm{eCap}} \tag{13c}$$

$$s_{2881}^{\text{cap-imp}}\leq 0.55U^{\mathrm{eCap}} \tag{13d}$$

$$f_{\ell,\tau_1+h}^{\mathrm{e}}=f_{\ell,\tau_2+h}^{\mathrm{e}}\qquad t_1,t_2\in\mathcal{T}^{\mathrm{g}}\text{ if }\Omega_{\tau_1}=\Omega_{\tau_2},h\in\{1,2,\ldots,24\} \tag{13e}$$

$$p_{n'ti}\leq U^{\mathrm{powCap}} \tag{13f}$$

$$\tag{13g}$$

Constraints (13a)-(13d) impose the energy capacity limits and equations. The details of energy capacity limits are provided in Table S12. The operations of the import node are carried out for the entire year, rather than over representative days like other power nodes. Therefore, in constraints (13e) we enforce the flow between the import node and other power nodes to be the same for any days that are represented by the same day. Finally, the constraints (13f) restrict the production capacity to its maximum limit.

8.3 Gas System Model

We now model the objective function and constraints pertaining to the gas system in the JPoNG.

Objective Function:

$$\min\sum_{l\in\mathcal{L}^{\mathrm{g}}}\left(C_\ell^{\mathrm{pipe}}z_\ell^{\mathrm{g}}+c_\ell^{\mathrm{pipeDec}}z^{\mathrm{gDec}}\ell+c_\ell^{\mathrm{pipeFix}}z_\ell^{\mathrm{gOp}}\right) \tag{14a}$$

$$+\sum_{k\in\mathcal{N}^{\mathrm{g}}}\sum_{\tau\in\mathcal{T}^{\mathrm{g}}}C^{\mathrm{ng}}g_{k\tau} \tag{14b}$$

$$+\sum_{j\in\mathcal{N}^s}(C_j^{\mathrm{strInv}}x_j^{\mathrm{gStr}}+C_j^{\mathrm{vprInv}}x_j^{\mathrm{vpr}}) \tag{14c}$$

$$+\sum_{j\in\mathcal{N}^s}\left(C^{\mathrm{strFix}}(I_j^{\mathrm{gStr}}+x_j^{\mathrm{gStr}})+C^{\mathrm{vprFix}}(I_j^{\mathrm{vpr}}+x_j^{\mathrm{vpr}})\right) \tag{14d}$$

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

$$+\sum_{k\in\mathcal{N}^{\mathrm{g}}}\sum_{\tau\in\mathcal{T}^{\mathrm{g}}}(C^{\mathrm{gShed}}a^{\mathrm{ng}}_{k\tau}) \tag{14e}$$

$$+\sum_{l\in\mathcal{L}^{\mathrm{LCF}}\setminus 1}C^{\mathrm{LCF}}_{l}(U^{\mathrm{LCF}}_{l}-U^{\mathrm{LCF}}_{l-1})\lambda^{\mathrm{LCF}}_{l} \tag{14f}$$

The objective function (14) minimizes the total investment and operating costs incurred in the gas system. The first term (14a) is the investment, strategic decommissioning and FOM costs for new, existing, and operational pipelines, respectively. The second term (14b) is the cost of procuring gas from various sources to the system. For example, New England procures its NG from Canada and its adjacent states such as New York. Term (14c) and (14d) handle the investment and FOM costs associated with gas storage, respectively. The term (14e) is the penalty for gas load shedding. The last term (14f) captures the cost of LCF consumption.

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

$$\begin{aligned}&g_{k\tau}-\sum_{l\in\mathcal{L}^{\mathrm{gExp}}_{k}}f^{\mathrm{g}}_{\ell\tau}+\sum_{l\in\mathcal{L}^{\mathrm{gImp}}_{k}}f^{\mathrm{g}}_{\ell\tau}-\sum_{n\in\mathcal{A}^{\mathrm{e}}_{k}}f^{\mathrm{ge}}_{kn\tau}\\&+\sum_{j\in\mathcal{A}^{s}_{k}}(f^{\mathrm{vg}}_{jk\tau}-f^{\mathrm{gl}}_{kj\tau})+a^{\mathrm{LCF}}_{k\tau}+a^{\mathrm{g}}_{k\tau}=D^{\mathrm{g}}_{k\tau}\end{aligned} \tag{15}$$

Constraints (15) state that for each node and period, the imported gas (i.e., injection), flow to other gas nodes, flow to power nodes, flow to and from storage nodes, load satisfied by LCF, and unsatisfied gas load should add up to demand. Unlike power flow, the flow in pipelines is assumed to be unidirectional, as is typical for most long-distance transmission pipelines involving booster compressor stations [15] , for which we ignore the relatively small electricity demand.

Flow on Representative Days:

$$f^{\mathrm{ge}}_{kn\tau_1}=f^{\mathrm{ge}}_{kn\tau_2}\qquad\qquad\tau_1,\tau_2\in\mathcal{T}^{\mathrm{g}}\text{ if }\Omega_{\tau_1}=\Omega_{\tau_2} \tag{16}$$

Given the set of representative days used to model power system operations (see Fig S12), constraint (16) ensures that gas consumption by the power system for all the days represented by the same day is identical.

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

$$L^{\mathrm{inj}}_{k}\le g_{k\tau}+a^{\mathrm{LCF}}_{k\tau}\le U^{\mathrm{inj}}_{k} \tag{17a}$$

$$\lambda^{\mathrm{LCF}}_{l}\le y^{\mathrm{LCF}}_{l}\qquad\qquad l\in\mathcal{L}^{\mathrm{LCF}} \tag{17b}$$

$$y^{\mathrm{LCF}}_{l}\le y^{\mathrm{LCF}}_{l-1}\qquad\qquad l\in\mathcal{L}^{\mathrm{LCF}}\setminus 1 \tag{17c}$$

$$\sum_{k\in\mathcal{N}^{\mathrm{g}}}\sum_{\tau\in\mathcal{T}^{\mathrm{g}}}a^{\mathrm{LCF}}_{k\tau}=\sum_{l\in\mathcal{L}^{\mathrm{LCF}}\setminus 1}\lambda^{\mathrm{LCF}}_{l}(U^{\mathrm{LCF}}_{l}-U^{\mathrm{LCF}}_{l-1}) \tag{17d}$$

The gas fuel import limits are imposed in constraints (17a). Constraints (17b) impose the share of consumption in the LCF availability at interval $[U^{\mathrm{LCF}}_{l-1},U^{\mathrm{LCF}}_{l}]$ is only positive if the consumption level exceeds U^{LCF}_{l-1}. Constraints (17c) ensures that the LCF consumption can not exceed level l before exceeding level $l-1$. The total consumption of LCF across all availability levels is calculated in constraints (17d).

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

$$f^{\mathrm{g}}_{\ell\tau}\le U^{\mathrm{pipe}}_{\ell}z^{\mathrm{gOp}}_{\ell} \tag{18a}$$

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

$$\sum_{k\in\mathcal{N}^{\mathrm{g}}:j\in\mathcal{A}_k^s} f_{kj\tau}^{\mathrm{gl}} = s_{j\tau}^{\mathrm{liq}} \tag{18b}$$

$$\sum_{k\in\mathcal{N}^{\mathrm{g}}:j\in\mathcal{A}_k^s} f_{jk\tau}^{\mathrm{vg}} = s_{j\tau}^{\mathrm{vpr}} \tag{18c}$$

The constraints (18a) limit the flow between gas nodes for operational pipelines, respectively. The flow to liquefaction facilities is calculated in constraints (18b). Similarly, the flow out of vaporization facilities is modeled via constraints (18c).

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

$$s_{j\tau}^{\mathrm{gStr}} = (1-\beta)s_{j,\tau-1}^{\mathrm{gStr}} + \gamma_j^{\mathrm{liqCh}} s_{j\tau}^{\mathrm{liq}} - \frac{s_{j\tau}^{\mathrm{vpr}}}{\gamma_j^{\mathrm{vprDis}}} \tag{19a}$$

$$s_{j\tau}^{\mathrm{vpr}} \le I_j^{\mathrm{vpr}} + x_j^{\mathrm{vpr}} \tag{19b}$$

$$s_{j\tau}^{\mathrm{gStr}} \le I_j^{\mathrm{gStr}} + x_j^{\mathrm{gStr}} \tag{19c}$$

Constraints (19a) ensure the storage balance. Constraints (19b) and (19c) limit the capacity of vaporization and storage tanks to their initial capacity plus the increased capacity, respectively.

Operational Pipelines:

$$z_\ell^{\mathrm{gOp}} = \mathcal{I}^{\mathrm{pipe}} + z_\ell^{\mathrm{gInv}} - z_\ell^{\mathrm{gDec}} \tag{20a}$$

Pipeline ℓ is operational if either it is existing and not decommissioned, or newly established.

8.4 Coupling Constraints

The following constraints are coupling constraints that relate operational decisions of the power and gas systems together.

$$\sum_{k\in\mathcal{A}_n^{\mathrm{e}}} f_{kn\tau}^{\mathrm{ge}} = \sum_{t\in\mathfrak{T}_\tau^{\mathrm{e}}} \sum_{i\in\mathcal{G}} h_i p_{nti} \qquad n \in \mathcal{N}^{\mathrm{e}}, \tau \in \mathfrak{R} \tag{21a}$$

$$\mathcal{E}^{\mathrm{e}} = \sum_{n\in\mathcal{N}^{\mathrm{e}}} \sum_{t\in\mathcal{T}^e} \sum_{i\in\mathcal{G}} w_t(1-\eta_i)\eta^{\mathrm{g}} h_i p_{nti}$$

$$\mathcal{E}^{\mathrm{g}} = \sum_{k\in\mathcal{N}^g} \sum_{\tau\in\mathcal{T}^g} \eta^{\mathrm{g}}(D_{k\tau}^{\mathrm{g}} - a_{k\tau}^{\mathrm{LCF}} - a_{k\tau}^{\mathrm{g}})$$

$$\mathcal{E}^{\mathrm{e}} + \mathcal{E}^{\mathrm{g}} \le (1-\zeta)(U_{\mathrm{emis}}^{\mathrm{e}} + U_{\mathrm{emis}}^{\mathrm{g}}) \tag{21b}$$

The first coupling constraints (21a) capture the flow of gas to the power network for each node and at each time period. The variable $\mathcal{E}^{\mathrm{e}}$ accounts for CO_2 emission due to the consumption of gas in the power system. The variable $\mathcal{E}^{\mathrm{g}}$ computes the emission from the gas system by subtracting the demand from LCF consumption and gas load shedding. The second coupling constraint (21b) ensures that the net CO_2 emissions associated with the power-gas system are below a pre-specified threshold value, which is defined based on a baseline (e.g., historical) emissions level. The first term is the emissions due to non-power gas consumption (i.e., gas consumption in the gas system such as space heating, industry use, and transportation). Since the model does not track whether LCF is used to meet non-power gas demand or for power generation, the first term computes gross emissions from all gas use presuming it is all fossil and then subtracts emissions benefits from using LCF.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Here we treat LCF as a carbon-neutral fuel source [32], and thus the combustion emissions associated with its end-use are equal to the emissions captured during its production. The second term captures the emission from NG-fired power plants. Alternatively, the emission constraints can be applied only to the power system as in [57] or applied separately to each system as in [15].

9 Supply-side Model Data

This section describes how we obtained parameters for both networks for the New England Case study. We start with the power system and explain the data preparation process. We then expound on the steps we took to construct the gas network and associated data inputs. In this study, our region is New England. However, the modeling framework and the solution approach are applicable to other regions to any spatial resolution.

Fig. S14 shows the name and boundaries of New England counties. The region currently has 67 counties with populations ranging from seven thousand to more than 1.6 million.

9.1 Power System Data

Our parameters for the power network are based on the *US Test System* developed by *Breakthrough Energy* [72]. The dataset contains high-temporal and spatial resolution load and VRE data for the entire US "base_grid" in the year 2016. It also contains load data for the year 2020 and projections for the year 2030. The dataset provides a test system with detailed information for existing buses, substations, plants, branches, and generation profiles for existing renewable energy capacity, including solar and land-based wind. We only use a subset of these datasets as explained in the following sections.

General Note about Tables: The first column of tables S7, S8, S10, S11, and S14 shows the associated symbol in the formulation. The notations with *tilde* are *crude* cost values whose manipulated forms (e.g., regional update, annualization) are used in the numerical model. For example, $\tilde{C}_i^{\text{inv}}$ in Table S8 denote the value of C_i^{inv} before annualization.

9.1.1 Power Nodes

The power network has 18 nodes, which, except for the import node 18, are located in New England regions and correspond to a group of counties with a more than 600,000 aggregate population. The locations of nodes inside New England are set to the centroid of the counties they represent. The number of nodes is selected to sufficiently capture the network effect without significantly compromising the problem scalability. All counties of a group are in the same state and form a contiguous landmass. Fig S15 shows the grouping of counties and the power node associated with each group. The import node 18 is located in the Quebec province of Canada.

9.1.2 Import Node

Roughly 14% of New England power consumption is imported [33], mostly from Quebec, and the amount is expected to significantly increase by 2050 [73]. Therefore, we define an import node in Quebec as a power system node, albeit with some

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

differences from other power nodes that are located inside New England. Node 18 in Fig. S15 shows the import node's location. Our modeling assumptions for the import node are based on [73]. The study represents the entire Quebec system as a single bathtub that combines all reservoir and run-of-the-river systems. This leads to a reservoir with a maximum energy capacity (U^{eCap}) of 175.5 TWh and a power capacity (U^{powCap}) of 41.1 GW. According to the study, the minimum stable output is 27% of Quebec's maximum power capacity. Also, it is assumed that the energy capacity of the reservoir should be at 70%, 55%, and 70% of its maximum capacity on January 1, May 1, and December 31, respectively. The ramp rate of power production is given at 14% of its maximum power capacity. The reservoir only incurs yearly FOM cost at $29312 per MW. Our model does not allow the investment of new power plants or storage facilities in the node.

There is currently a high-voltage transmission line between Quebec and a location in central Massachusetts which we assumed to be node 6 [73]. The maximum flow capacity of this line is 2 GW. There is another transmission line project with a capacity 1.2 GW under construction between Quebec and southwest Maine [74]. We assumed that this line is connected to node 9. The load and inflow rates for 2050 are provided in [73]. Around 90% of Quebec's power load is already satisfied by hydropower and the rest mainly by solar and onshore wind generation. Since we only consider hydropower generation, we assume the power load of node 18 to be 90% of its projected load for 2050.

9.1.3 Transmission Lines

In the *Breakthrough Energy* [72] data, we consider the base_grid and start by filtering the data for New England region, which corresponds to zones 1 to 6 in the dataset. We then filter for high voltage buses [75] i.e., those with voltage greater than or equal to 345kV. This process results in 188 nodes, each of which represents a high-voltage bus. The filtering process results in 192 transmission lines with known susceptance and maximum flow limit. We use this data to identify the existing lines between power nodes. We first assign each of the 188 buses to its nearest county, and then keep the existing lines between any two counties belonging to different power nodes. This process resulted in 30 existing transmission lines, with total capacity of 103 GW, in the power network.

Once the existing transmission lines are identified, we create candidate lines. We assume that each node can be connected to the two nearest nodes via candidate transmission lines, thus creating 34 candidate lines. The susceptance and maximum flow limit of these lines is set to their average in the set of existing lines. Fig. S16 shows all the existing and candidate transmission lines in the current data set. It is worth noting that multiple transmission lines connect some node pairs, but only one is depicted in the figure.

9.1.4 Power Plants

The "base_grid" data contains power plant information at each bus. We first remove power plant types of "dfo" (distillate fuel oil) and "coal" as their share is not substantial in the generation fleet of the historical year (i.e., 2016), and there are plans to completely phase out those plant types by 2050. We then calculate the existing generation capacity at each bus from each plant type by considering "in-service" plants with generation capacity greater than 10MW for "ng" (i.e., gas-fired plants) and

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

"nuclear" type and greater than 2MW for VREs. Note that "ng" plants are treated as a lump since the Breakthrough Energy data set does not provide a breakdown between the capacity of combined-cycle and open-cycle plants. We then assign each plant to the nearest node and eventually aggregate the generation capacity of each plant type for each of the 17 nodes. Tables S7 and S8 present the technical assumptions for the existing and new power plants used in this study. Table S6 presents the nameplate capacities for various generation technologies at each power node. Most of the parameters for the existing plants are derived from the Breakthrough Energy data set [76]. The footnote text in Table S7 presents details on the value of each parameter. The technical assumptions for new plants are largely derived from National Renewable Energy Laboratories (NREL) Annual Technology Baseline 2021 edition (ATB 2021) for the year 2045 [56]. Values of parameters not available in ATB 2021 are adopted from the corresponding existing power plants or obtained from Sepulveda et al. [57]. Details of each parameter are provided in the footnote text in Table S8.

Table S6: Aggregate nameplate capacity of existing plants at each node (MW)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
ng	1327	80	3549	1807	23	848	630	805	0	1024	1526	1526	1811	1359	1235	959	
solar	25	58	9	46	102	154	75	0	0	57	0	0	10	0	0	15	10
wind	2	0	2	0	8	3	56	686	213	115	0	185	51	0	0	5	0
hydro	0	38	0	0	0	4	1822	406	199	280	33	219	0	68	48	8	0
nuclear	1226	0	0	0	617	0	0	0	0	0	0	0	0	0	0	0	1889

Table S7: Parameter Values for Existing Plants

Symbol	Parameter\Type	gas	solar	wind	hydro	nuclear
	FOM [$/kW/year][1]	21	23	43	78	145
C_i^{var}	VOM [$/MWh][2]	5	0	0	0	2
η_i	CO_2 Capture [%][3]	0	-	-	-	-
h_i	Heat Rate [MMBtu/MWh][4]	8.7	0	0	0	10.6
$\tilde{C}_i^{\text{dec}}$	Decom. cost per plant [$][5]	5.0e6	4.5e4	1e6	-	3.0e8
U_i^{prod}	Nameplate capacity [MW][6]	137	4.7	47	15	933
L_i^{prod}	Min stable output [%][7]	31	0	0	0	42
U_i^{ramp}	Hourly Ramp rate [%][8]	96	-	-	-	25
C_i^{fix}	FOM per plant [$/MW/year][9]	3.6e6	1.45e5	1.8e6	1.8e6	1.4e8
C_i^{startUp}	Startup Cost [$][11]	4.52e4	-	-	-	4.6e4

[1] and [2] from [56] in year 2019, [3] from [77], [4] approximated from linear and quadratic coefficient of heat rate curve provided in [76], [5] estimated from [78] except for nuclear which is obtained from [79]. Decommissioning of hydro plants is not considered, [6] and [7] approximated from [76], [8] estimated from 30-min ramp rate in [76], [9] $r1 * 1000 * r6$ where $r1$ is the FOM value in $r1$ and $r6$ is nameplate capacity provide in row6, [11] from [57]. Startup cost for "ng" type is assumed to be the same as "CCGT" type.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Table S8: Parameter Values for New Plants

Symbol	Parameter\Type	OCGT	CCGT	CCGT -CCS	solar -new	wind -new	wind-off. -new	nuclear -new
	CAPEX [$/kW][1]	780	935	2167	672	808	2043	6152
	FOM [$/kW/year][2]	21	27	65	15	35	74	145
C_i^{var}	VOM [$/MWh][3]	5	2	6	0	0	0	2
η_i	CO_2 Capture [%][4]	0	0	90	-	-	-	-
h_i	Heat Rate [MMBtu/MWh][5]	9.72	6.36	7.16	0	0	0	10.46
	Lifetime [year]	30	30	30	30	30	30	30
U_i^{prod}	Nameplate capacity [MW][6]	237	573	400	10	10	10	360
C_i^{inv}	CAPEX per Plant [$][7]	1.85e8	5.36e8	8.67e8	6.72e6	8.01e6	2.04e7	2.21e9
L_i^{prod}	Minimum stable output [%][8]	25	33	50	0	0	0	50
U_i^{ramp}	Hourly ramping rate [%][9]	100	100	100	-	-	-	25
C_i^{fix}	FOM per plant [$/yr][10]	5.0e6	1.55e7	2.6e7	1.5e5	3.5e5	5.22e7	5.22e7
C_i^{startUp}	Startup Cost [$][11]	8.0e3	4.52e4	3.79e4	-	-	-	4.6e4

[1−5] from [56] in year 2045. For CCGT-CCS, "Conservative" technology class is considered. For "wind-new" and wind-offshore "Moderate-Class4" technology class is considered. For all others "Moderate" cost assumption is considered, [6] from [57] for "OCGT", "CCGT", CCGT-CCS, and "nuclear-new". For VRE, a modular capacity of 10MW is considered, [7] $r1 * 1000 * r6$ where $r1$ is the CAPEX value in row[1] and $r6$ is nameplate capacity provide in row[6], [8] from [57], [9] from [57], [10] $r2 * 1000 * r6$ where $r2$ is the FOM value in $r1$ and $r6$ is nameplate value provide in $r7$, [11] from [57]

9.1.5 Offshore Wind Siting

The development of large-scale offshore wind power generation is currently ongoing only for Massachusetts and Rhode Island [80], hence we allow the establishment of offshore wind plants only at nodes 1, 2, 3, 4, 5, and 13.

9.1.6 Power Load

We obtain the residential load for each of the electrification scenarios from the bottom-up approach introduced in the paper and elaborated further in Section 12.1. To account for non-residential load, we consider the 2050 hourly load profiles provided as part of the NREL's Electrification Future Study (EFS) Load Profile dataset [81]. The repository contains hourly load profiles for various electrification (Reference, Medium, High) and technology advancement (Slow, Moderate, Rapid) scenarios. The load profiles are provided for several years, including 2050, and are further disaggregated by state, sector (residential, commercial, etc.), and subsectors (space heating and cooling, water heating, etc.). For all scenarios, we consider the aggregated state-level hourly demand profile for the *High* electrification level with *Moderate* technology advancement. The state-level non-residential load is further disaggregated to the power node level based on the population share of each node throughout the entire state.

9.1.7 Regional Cost Multipliers

We incorporate regional capital cost multipliers provided in *ReEDS Model Documentation* [75], summarized in Table S9, to distinguish the capital costs of new power plants in the different model regions. These multipliers are applied to the baseline capital costs reported in Table S8. Subsequently, the capital costs are annualized to be included in the single-stage investment planning model using the following formula

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

for the annual cost fraction: $\frac{\omega}{1-(\frac{1}{1+\omega})^{lt}}$. Here, lt is the lifetime of the specific technology and ω corresponds to the discount rate of 7.1%. Thus, the annualized CAPEX for new power plants is obtained by multiplying the CAPEX by the annual cost fraction and regional multiplier. For every other investment cost (i.e., transmission lines, pipelines, storage, etc.), no regional cost multiplier is considered and we only multiply the CAPEX by the annual fraction factor to get the annualized CAPEX.

Table S9: Regional CAPEX multipliers for new plant types

State/Technology	OCGT	CCGT	CCGT -CCS	solar -new	wind -new	wind-off. -new	nuclear -new
Connecticut (CT)	1.25	1.3	1.3	1.15	1.4	1.1	1.1
Massachusetts (MA)	1.1	1.1	1.1	1.05	1.35	1.1	1.05
Maine (ME)	1.25	1.3	1.3	1.1	1.35	1.1	1.1
New Hampshire (NH)	1.1	1.1	1.1	1.05	1.35	1.1	1.05
Rhode Island (RI)	1.2	1.25	1.25	1.1	1.35	1.1	1.05
Vermont (VT)	1.1	1.1	1.1	1.05	1.35	1.1	1.05

9.1.8 Plant Decommissioning Cost

The decommissioning costs for power plants and pipelines are not annualized. Instead, we assume a gradual retirement process starting from the mid-year date of 2040 until the planning year 2050. Since the model only considers a time horizon of a single year, we divide the decommissioning costs for these assets by 10. The distribution of the decommissioning cost over multiple years allows the model to decommission the asset if it is not utilized and its fixed cost is higher than the distributed decommissioning cost. Otherwise, the model keeps the idle asset without decommissioning it.

9.1.9 Power Storage

Energy storage is likely to be an essential part of the future power systems dominated by VRE supply. While many storage technologies are proposed or under development, we model Li-ion batteries. The technology provides short-term storage, usually around four hours. We sourced Li-ion battery cost and performance assumptions from [56] as summarized in Table S10.

9.1.10 Reserve Margin

The capacity derating factor $\gamma_{nit}^{\mathrm{CRM}}$ for VRE plants is set to their capacity factors. For other plants, the value is set to 1. The capacity reserve margin rate R^{CRM} is set at 15%, which is the value recommended by North American Electric Reliability Corporation (NERC) [83].

9.1.11 Transmission Line FOM

The fixed cost for transmission lines is estimated from [84], in which the operating and maintenance cost for a transmission line of 134 kV is given as about 20% of the

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Table S10: power storage parameters

Symbol	Parameter	Li-ion
$\tilde{C}^{\text{EnInv}}$	Energy capital cost [$/kW]	
$\tilde{C}^{\text{pInv}}$	Energy power cost [$ /kWh]	156[4]
γ_r^{eCh}	Charge efficiency	0.92[5]
γ_r^{eDis}	Discharge efficiency	0.92[6]
C_r^{EnInv}	Energy related FOM ($/kWh/year)	3.22[7]
C_r^{pInv}	Power related FOM ($/kW/year)	3.9[8]
γ_r^{selfD}	hourly self-discharge rate	2.08e-5[9]
	Lifetime	15[10]

columns [1] and [2] from [31],
[3] and [4] from [56] in year 2045 averaged over "Advance", "Moderate" and "Conservative" estimates, [5] and [6] from [56] where the round-trip efficiency is provided at 85%, [7] and [8] 2.5% of energy capital and power cost (row 1 and 2), respectively [82], [9] from [31] the monthly self-discharge provided at 1.5%, [8] from [56]

investment cost. Therefore, we estimate the FOM cost as 20% of the CAPEX over its lifetime. We currently consider CAPEX of a line at 7398 $/MW/mile over 30 years so that the annual FOM will be $C_\ell^{\text{trFix}} \frac{7398 \times 0.2}{30} \approx 49.3$$/MW/mile for all $\ell \in \mathcal{L}^{\text{e}}$. Investment and FOM costs of a line can be calculated by multiplying these values by the length and maximum capacity of a line.

9.1.12 CCS parameters

We based our estimation of CCS parameters on [85–87]. We assume that the collected carbon is stored in Appalachian Basin at a Marcellus region located in the middle of Pennsylvania [86]. We then calculate the distance between each node and the storage site. The total capacity of the Appalachian Basin is 1278 Mt (megaton). Assuming the Basin is operable for 100 years, the total annual carbon storage capacity U^{CCS} becomes 12.78 Mt. Other parameters are calculated as follows:

- $C_{\text{CO2}}^{\text{inv}}$: Reference [87] provides CAPEX for 10 and 100 miles pipelines. We consider 100-mile pipelines as all distance values are greater than 200 miles. The CAPEX and FOM for 100 miles are 225 $M and 1.3 $M (million dollars), respectively. With 30 years of a lifetime for CO_2 pipelines and WACC=7.1%, the CAPEX becomes 18.31 $M. The FOM is given at 1.3 $M, so the per mile investment and FOM is (18.31e6+1.3e6)/100 = 196e3 $/mile. The reference assumes that the capacity of the pipeline is 10 Mt/y (megaton per year). Therefore, the levelized investment and FOM becomes 196e3/10e6 = 0.0196 $/mile/ton.
- $C_{\text{CO2}}^{\text{str}}$: The Appalachian basin is an aquifer type storage. The CAPEX is given at 4.3 $M in [87] for a storage site of type aquifer with 7.3 Mt capacity per year. The CAPEX consists of injection site screening and evaluation, injection equipment, and drilling of 6 wells. Assuming 30 years of lifetime, the annualized CAPEX becomes 350e3 $. The FOM is given at 600e3 $, so the levelized investment and FOM is (350e3+600e3)/7.3e6 = 0.13 $/ton.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

- E^{pipe}: For a pipeline of 100 miles long with a capacity 10 Mt/year, reference [87] gives the electric requirement at 32,000 MWh/year or equivalently 32/8760e6 = 0.00365e-6 [MWh/mile/ton/hour].
- E^{pipe}: For a pipeline with a capacity of 10 Mt/year, each pump consumes 4190 MWh electricity annually which accounts for 4190/(8760e6)=0.478e-6 MWh/ton/hour.
- E^{cprs}: Compression pump are located every 3.3 miles along the pipeline [86]. Therefore, the value is $d_n/3.3$, where d_n is the distance of the node from the carbon storage location in miles.

We assume that CAPEX for compression pumps is negligible. Also note that these cost estimations for CCS storage can be an underestimation as we levelized the investment and FOM costs and do not consider other cost parameters such as labor, FOM of compression pumps, and fugitive emission amount which are listed in [86].

9.1.13 Other Parameters

Other economic and technical assumptions for the power network are presented in Table S11 and Table S12.

Table S11: Other parameters for power network

Symbol	Parameter	Value
	Weighted average cost of capital (WACC)[1]	7.1%
$\tilde{C}_{\ell}^{\text{trans}}$	Transmission line investment cost [$/MW/mile][2]	7398
	Li-ion Battery lifetime [year][3]	30
	Transmission line lifetime [year][4]	30
	Uranium price [$/MMBtu][5]	0.72

[1] from [57], [2] average of costs reported in Table 9 of [73], [3] from [56], [4] from [56] in year 2045, [5] from [57] in year 2045

Table S12: Resource Availability Data

$\mathcal{Q}$	Maximum Available Value
"solar", solar-UPV	22 GW
"wind", "wind-new"	10 GW
wind-offshore	280 GW
"nuclear", "nuclear-new"	3.5 GW
"hydro"	2.6 GW

Resource availability amounts are obtained from [29] except for "hydro" which is obtained from [73] such that 853 MW is run-of-river and 1768 MW is pumped storage hydro power

9.2 Gas System Data

The gas network consists of two types of nodes: 1) Gas nodes that have injection capacity and load and 2) SVL nodes. We further distinguish between gas nodes whose

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

load is zero but can be used to inject gas in the system and those that have gas load. The nodes without load are created in the boundary of the region to represent the connection of gas nodes to the outside region. We call the former set of nodes "boundary nodes" and the latter "load nodes." Note that this distinction is for the exposition of the data input and has no modeling implications. Each SVL node consists of Storage, Vaporization, and Liquefaction (SVL) facilities, each with its own capacity. This section provides details of each node type and other parameters associated with the gas system.

9.2.1 NG Nodes

We construct the gas network based on the data available on the Energy Information Administration (EIA) website [58]. The pipeline data provides information on the interstate and major pipelines, and their start and ending counties. Each pipeline is a uni-directional means of transferring gas with a daily capacity limit. We first filter all the pipelines whose ending counties are one of the New England counties. We consider each of these counties as a load node. Some load nodes are connected to regions outside New England via pipelines. To capture the import of gas to the region via pipelines, we create boundary nodes in locations where a pipeline connects a load node to an outside region. This process resulted in six boundary nodes and 17 load nodes. The import of gas fuel to the region is realized either through pipelines from neighboring regions or coastal import facilities from overseas [68, 88]. We consider the boundary nodes as ground injections. The injection to the system can also be realized by LNG facilities located in Everett and Cape Ann, Massachusetts [68] that are equipped with vaporization plants. Our model does not differentiate between LNG and NG in terms of cost, consumption, or any other characteristics.

NG is usually stored in its liquefied form called LNG. LNG storage usually involves three types of facilities:

- Storage tank: smaller tanks located inland that are filled by truck supply from the import facilities and are used to regulate the pressure of the gas system – these facilities have tanks and vaporization facilities [89]. The capacity of a storage facility is measured in energy/volume level (i.e., MMBtu or MMScf)
- Liquefaction: a facility that receives NG from a pipeline system and liquefies it at a temperature of around -162 Celsius [89, 90]. The capacity of the liquefaction facility is measured in terms of flow rate (i.e., MMBtu/period or MMScf/period)
- Vaporization: The facility evaporates the LNG leaving the storage facility by warming it with seawater or air to produce gas that is injected into the pipeline network. In certain areas, trucks are used to transport LNG from a major storage site to a smaller storage site where they are later evaporated [89]. For ease of transportation, storage and vaporization facilities are built at the same location. The capacity of the vaporization facility is measured in terms of flow rate (i.e., MMBtu/day or MMScf/day).

We construct the SVL network based on the data provided in [68]. We assume that all three facilities are located at the same location. There are currently five liquefaction and 43 storage facilities in New England. The exact location of these

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

facilities is not provided, but the source provides a map of the region showing the approximate locations of storage tanks. Given that there are only five liquefaction facilities, we first cluster storage tanks into five locations and assume a liquefaction and a vaporization facility at each location. This assumption effectively approximates the practice of moving LNG via trucks from the centralized liquefaction facilities to the distributed storage (and vaporization) facilities. Each of these locations is a node in the SVL network. The total liquefaction, vaporization and storage capacities for the New England region are given in [68]. To account for variations in the capacities, we unevenly distribute these capacities over five SVL. This process resulted in six boundary nodes, 17 load nodes, and five SVL nodes as depicted in Fig. S17. Note that node 9 also

9.2.2 Pipelines

We use the pipeline data from [88] to identify the existing pipelines between load nodes and between boundary nodes and load nodes. Each existing pipeline has a known daily capacity limit. We construct candidate lines for boundary nodes by creating a pipeline between them to three of their nearest load nodes.

The transfer of gas between different nodes is assumed to be realized as follows:

- Boundary and load nodes to load nodes: The existing interstate pipelines between boundary and load nodes are provided in [88]. We estimate the existing intrastate pipelines based on the map provided in S18 with capacities associated with the interstate pipeline ending in the node from which the intrastate pipeline originates. For candidate pipelines, we assume that each node is connected to 2 nearest non-adjacent, non-boundary load nodes, resulting in 46 candidate pipelines. The capacity of candidate pipelines is set to the average capacity of the existing ones.
- Gas load node to power node: The location of gas-fired power plants as well as gas pipelines is provided in [91]. Fig. S18 shows that each power plant is connected to one major pipeline. We use this fact to assume that each power node is already connected to its nearest NG node through distribution pipelines. The pipeline capacity available to these neighboring gas nodes thus limits the amount of flow between gas and power nodes, hence we do not consider a limit to pipelines connecting gas and power nodes.
- Gas load nodes and SVL nodes: We assume that each gas load node is already connected to its nearest SVL node through two sets of pipelines. The first set carries gas from the gas node to the liquefaction facility in an SVL, and the second set of pipelines transports gas from the vaporization facility in an SVL to a gas node.

Fig. S19 illustrates the existing pipelines of all node types of the gas system. Currently, in the year 2022, there are 25 pipelines that connect boundary and load nodes. The procedure that we considered to create candidate lines resulted in 46 potential connections that are depicted in Fig. S20. The power nodes draw gas from their nearest load nodes as illustrated in Fig. S21. Finally, Fig. S22 shows all the connections we consider in JPoNG.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

9.3 LCF Availability

The supply curve for LCF modeled in this study is based on limited available analysis in the literature on the potential cost and availability of LCF for the Northeast region. In developing the LCF supply curve, we primarily focused on supply and cost of bio-methane or renewable natural gas.

We start by estimating the availability of biogenic LCF in New England. The annual production potential of New England for 2040 is given in [92] for three scenarios, namely low, high, and technical resource potential. We use the technical resource potential in 2040 reported by this study, at 250.8 tBty [92]. The other source of LCF to be used in New England is the production in other states as well as the neighboring regions. New York is a likely neighboring state that has plans for biogenic LCF and may be able to partially provide for New England's LCF needs. We used the maximum production potential of biogenic LCF in New York State in 2040, estimated by another study at 272.3 tBTu under a maximum growth scenario [93]. We further assume that consumption of biogenic LCF among subregions in the US Northeast (New England and New York) is proportional to their current natural gas consumption. In 2022, New England's share was 41% of the total natural gas consumed in the US Northeast [88]. Accordingly, the combined biogenic LCF potential for the US Northeast is 250.8+272.3=523.1 tBtu, of which 0.41*523.1=214.5 tBtu is assumed to be available for consumption in New England.

The biogenic supply of LCF is priced based on a recent study on LCF supply for Massachusetts that estimated that the first 33% of biogenic LCF could be produced at 10$/MMBtu while the rest would be at around 25$/MMBtu [94]. Collectively, the above biogenic supply of 214.5 TBtu and the two price levels define the first two bins of the supply curve as shown in Table S13.

Any amount of LCF in excess of the projected biogenic supply of 214.5 tBtu is assumed to be sourced from synthetic methane production, which involves coupling biogenic CO2 sources with renewable H2. This process will require electricity input that was not explicitly considered in our analysis. As a conservative estimate, we priced the LCF supply above 214.5 TBtu at 50$/MMBtu. Interestingly, across the modeled high electrification scenarios (HE, HX), we find that LCF consumption stands at 70-277 TBtu indicating limited to no reliance on synthetic methane supply under high electrification of buildings sector.

Availability Intervals (tBtu)	[0, 70.8]	(70.8, 214.5]	(214.5, 1000]
Price ($/MMBtu)	10	25	50

Table S13: LCF availability and price levels. About 33% of total capacity is produced at 10 $/MMBtu, and the rest at 25$/MMBtu. Any amount beyond that, is assumed to be produced by synthetic fuel at a price 50$/MMBtu. The last interval is capped by 1000 tBu as a large upper bound for modeling purposes.

9.3.1 Pipelines FOM Costs

The FOM cost of gas pipelines is estimated from [95] in which it is stated that 5-10% (on average 7.5%) of a pipeline's cost across its lifetime is operating cost, of which 72% is the fixed cost. Since we assume a 50-year lifetime for new pipelines [96], the

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Table S14: Other parameters for NG/SVL network

Parameter	Value	
$\tilde{C}_j^{\text{strInv}}$	Storage tank CAPEX[1] [$/MMBtu]	729.1
$\tilde{C}_j^{\text{vprInv}}$	Vaporization CAPEX[2] [$/MMBtu]	1818.31
C_j^{strFix}	Storage tank FOM[3] [$/MMBtu]	3.6
C_j^{vprFix}	Vaporization FOM[4] [$/MMBtu/d]	327.3
γ_j^{ligCh}	Liquefaction charge efficiency (%)	100
γ_j^{vprDis}	Vaporization discharge efficiency[5] (%)	98.9
C_j^{gShed}	Gas load shedding cost [$/MMBtu]	1000
η^{g}	Emission factor for NG [6] [ton/MMBtu]	0.053
C^{alt}	LCF price [$/MMBtu][7]	20
	SVL lifetime [year]	30
$\tilde{C}_\ell^{\text{pipe}}$	Pipeline investment cost [$/mile][8]	20e6
C^{ng}	NG price [$/MMBtu][9]	5.45
	Pipeline lifetime[10] [year]	50

[1] from Table 1 of [98], [2] from Table 1 of [98], [3] 0.5% of CAPEX according to [90], [4] 1.8% of CAPEX according to [90], [5] estimated from [89], [6] from [99], [7] from [32], [8] approximated from "Pipeline projects" provided in [88] for gas pipeline projects. We consider the completed new pipeline projects in New England between 2011 to 2020 and divide the total cost to the length of the pipeline, [9] from [57] [10] from [96]

FOM cost for pipelines is calculated as $C_\ell^{\text{pipeFix}} = \frac{20e6 \times 0.075 * 0.72}{50} = 21600$/mile for all $\ell \in \mathcal{L}^{\text{g}}$.

9.3.2 Pipelines Decommissioning Cost

A cost estimate for decommissioning of pipelines in the US Gulf of Mexico between 1995 to 2015 is provided in [97] is provided at around 3e5 $/mile for offshore pipelines. However, we could not find any substantial study estimating the decommissioning cost for the onshore pipelines. Furthermore, unlike offshore pipelines, onshore pipelines can be repurposed to be used by other energy carriers such as hydrogen. Therefore, it is not clear whether the decommissioning of pipelines will carry any cost at the end of their lifetime value, hence we set the decommissioning cost to zero for all pipelines.

9.3.3 Other Parameters

Other economical and technical assumptions for the gas system are presented in Table S14.

9.3.4 NG Load

Similar to the power system, the gas load for the residential sector is obtained from the bottom-up approach proposed in this paper. The load for gas is generally disaggregated into five sectors, including residential, commercial, industrial, vehicle fuel consumption, and electric power customers [88]. The gas consumption for electric power generators is a decision variable in our model, so our input for gas demand

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

involves the gas demand in the remaining three sectors. The monthly state-level gas load for all five sectors is available on the EIA website, "Natural Gas Data" page [88]. The EIA website provides no information on the distribution of the monthly demand over its days (i.e., the load shapes). Therefore, we consider industrial, commercial, and vehicle fuel consumption in 2017 and uniformly distribute the monthly load across their corresponding days. We then scale the values based on the annual industrial and vehicle fuel consumption in 2050 under reference and high electrification scenarios. Finally, we aggregate loads for all sectors and subsectors to obtain the gas consumption of each state in 2050 under the two electrification scenarios. Once the daily gas demand is obtained for each state, we disaggregate the demand over each node in the proportion of the population each node represents in 2019 [100], which implicitly assumes that there is no change in the proportion of the population by 2050.

9.4 Emission Amounts

All New England states have set a goal to reduce the emission of GHG in 2050 by at least 80% below a baseline year, which is 1990 for all states except Connecticut [61]. The 80% goal has been subject to discussion in recent years to increase, with Vermont already setting the target at 95%. The total CO_2 emission for the New England states was 171.2 metric tons (mt) in 1990, of which 43.9mt was electricity energy-related emission and 23.6mt was NG energy-related emission [101]. The remaining emission was caused by consuming coal and petroleum that we do not consider in this study. Based on these figures, $U^{\mathrm{g}}_{\mathrm{emis}} = 23.6\mathrm{e}6$ and $U^{\mathrm{e}}_{\mathrm{emis}} = 43.9\mathrm{e}6$.

10 Capacity factor calculations

In our capacity factor modeling, we use weather data most fit-for-purpose for modeling the individual resource types, primarily using inputs specifically created for and calibrated to the usage in production modeling. We do not include climate change impacts in the capacity factor modeling.

10.1 Solar capacity factors

We calculate solar capacity factors for the 2001-2020 weather years using the pvlib library in Python [102]. We leverage solar data from the National Solar Radiation Database, which is more fit for purpose than the ERA5 data used in our demand-side analysis. Our input parameters follow the assumptions of Brown & O'Sullivan [103], except we do not model inverter efficiency losses (which are already considered in our cost parameters). We use the default incidence angle modifier assumptions in the pvlib physical model and assume an anti-reflective coating index of 1.3. Capacity factors are calculated for each weather location used in the demand-side analysis. Due to the cold climate of the study region, we use NREL's snow coverage model to approximate the hourly impact of snow coverage on solar capacity factors. We use snow accumulation data from the ERA5 data used in the demand analysis. The data is provided in millimeters of water equivalent, which we convert to inches of snow accumulation by using the one-equation model presented in [104]. While the one-equation model is shown to have a negative bias, it is useful as a rough approximation in our analysis, particularly because snow depth is only one component of the model for snow-related capacity derating. The snow coverage model is described in [105],

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

which calculates the fraction of the panel covered in snow given the hourly snowfall, outdoor temperatures, panel geometry, and an empirically defined coefficient describing the rate that snow slides off the panel. We assume the capacity derating is proportional to the fraction of panel surface covered by snow. Because the model is specific to fixed-tilt systems and we model systems with single-axis tracking, we approximate the slide rate for a panel with a tilt equal to the latitude of the panel site. We do not consider the likelihood that climate change reduces the frequency and intensity of snowfall events.

Capacity factors for each zone are calculated by averaging the capacity factors for the sites located within them. We use the same sites as in the demand analysis, shown in Fig. S28. For zones that do not include a site, capacity factors are taken from the site closest to the zone's centroid. Our solar capacity factors averaged 20.3% across the zones for 20 years. This is in agreement with NREL estimations of capacity factors in the region for 2019 [106].

10.2 Onshore and offshore wind capacity factors

We calculate wind capacity factors for the 2001-2020 weather years using publicly-available wind speed data provided by ISO-NE [107]. The data is provided for a set of existing and hypothetical onshore and offshore plant locations, shown in Fig. S23.

For each site, we calculate wind capacity factors. We use power curves to convert wind speeds into power output. For our modeled wind plants, we assume an onshore hub height of 100 m and an offshore hub height of 150 m. The ISO-NE wind dataset includes wind speeds at differing heights for each location which often do not match these standard heights. We adjust the speeds according to power-law approximations of the wind profile. The typical exponent value used in the onshore wind modeling literature is 1/7. For offshore, we assume an exponent of 0.11 [108]. We note that these adjustments generally only change the speeds by 1-2%. We calculate our capacity factors as in Brown and Botterud [34]. However, because we lack pressure data at hub height, we assume the air density at hub height is approximately equal to the standard air density used in the power curves. The onshore power curve is borrowed from the "Gamesa: G126/2500" turbine used in Brown and Botterud and the offshore power curve is taken from [109]. Both curves are shown in Fig. S24 below. In addition to the high-speed cutoffs shown in the power curves, we assume that the capacity factor for a given site is zero if the temperature at hub height is less than -30 C. To approximate hourly temperature at hub height, we use surface temperature data from the ERA5 dataset used in the demand-side analysis and then assume a decrease of 6.5^oC per 1000m of altitude gain, a lapse rate common in atmospheric science. Because the temperature data does not reflect the effects of climate change, which will likely result in warmer temperatures at hub height and less frequent cut-offs, this method is conservative. We uniformly reduce the capacity factors by 19% to reflect other system losses and downtime in alignment with Brown and Botterud [34].

We assign capacity factors for onshore wind to each zone in a manner analogous to the solar capacity factors, either averaging the capacity factors of the sites located within the zone or taking data from the site closest to the zone. Offshore wind sites are assigned to each state based on the state associated with the site in the ISO-NE dataset, which generally corresponds to the states granted the relevant BOEM leases or, for hypothetical sites, the state closest to the site. We only calculate offshore wind capacity factors for Massachusetts and Rhode Island, as they are the only states with existing BOEM leases at the time of our analysis and therefore the most

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

likely to have significant operating offshore generation in 2050. Each coastal zone within a given state is assumed to have the same offshore wind capacity factors. The non-coastal zones do not allow for offshore wind generation as a candidate resource. Our average onshore wind capacity factors across the 20 years vary from 24.0% to 49.5% depending on the zone, with an average of 39.3%. Our offshore wind capacity factors average 49.8% across the states. The values are generally in alignment with expectations for these technologies [56, 110].

11 Computational Experiment Configurations and Solution Approach

The JPoNG code is developed in Python using the Gurobi solver. The details of the code and implementation are given in the subsequent section. The code along with all the data used in this study are available in the GitHub repository [111]. All instances are run on the MIT Supercloud system which uses an Intel Xeon Platinum 8260 processor with up to 48 cores and 192 GB of RAM [112]. The system currently uses Gurobi 10.1.

Our preliminary runs showed that for the given power-gas network, the resulting JPoNGbecomes a large-scale mixed-integer linear program (MILP), thus intractable within a reasonable time frame and for a sufficient number of representative days. Therefore, we propose a four-step approach (FSA) heuristic to approximate the optimal solution of instances. The heuristic utilizes the fact that relaxing certain constraints and variables can potentially reduce computational burden significantly without deviating much from the optimal solution. The steps of FSA are as follows:

1. Apply the following simplifications and solve the problem and get the solution $\mathcal{X}^1$
   - Relax integrality of all integer variables
   - Remove unit commitment variables and constraints
   - Remove any network variables and constraints
2. Apply the following simplifications and constraints and solve the problem and get the solution $\mathcal{X}^2$
   - Relax integrality of all integer variables
   - Remove any network variables and constraints
   - Add constraint $x_{ni}^{\text{op}} \geq x_{ni}^{\text{op}^1}, n \in \mathcal{N}^{\text{e}}, i \in \mathcal{H}$
   - Add constraint $z_{\ell}^{\text{gOp}} = \lfloor z_{\ell}^{\text{gOp}^1} \rceil, \ell \in \mathcal{L}^{\text{g}}$, where $\lfloor . \rceil$ rounds the value to the nearest integer.
3. Apply the following simplifications and constraints and solve the problem and get the solution $\mathcal{X}^3$
   - Relax integrality of all integer variables
   - Remove unit commitment variables and constraints
   - Add constraint $x_{ni}^{\text{op}} = \lfloor x_{ni}^{\text{op}^2} \rceil, n \in \mathcal{N}^{\text{e}}, i \in \mathcal{H}$
   - Add constraint $x_{ni}^{\text{op}} \geq \lfloor x_{ni}^{\text{op}^2} \rceil, n \in \mathcal{N}^{\text{e}}, i \in \mathcal{P}, i \in \mathcal{N}^{plt}$ where $\mathcal{N}^{plt}$ denote the set of new power plant types

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

- Add constraint $z_{\ell}^{\mathrm{gOp}} = z_{\ell}^{\mathrm{gOp}^2}, \ell \in \mathcal{L}^{\mathrm{g}}$

4. Apply the following simplifications and constraints and solve the problem and get the feasible solution $\mathcal{X}^4$
   - Relax integrality of all integer variables
   - Add constraint $x_{ni}^{\mathrm{op}} = \lfloor x_{ni}^{\mathrm{op}^2} \rceil, n \in \mathcal{N}^{\mathrm{e}}, i \in \mathcal{H}$
   - Add constraint $x_{ni}^{\mathrm{op}} \geq \lfloor x_{ni}^{\mathrm{op}^2} \rceil, n \in \mathcal{N}^{\mathrm{e}}, i \in \mathcal{P}, i \in \mathcal{N}^{plt}$ where $\mathcal{N}^{plt}$ denote the set of new power plant types
   - Add constraint $z_{\ell}^{\mathrm{gOp}} = z_{\ell}^{\mathrm{gOp}^2}, \ell \in \mathcal{L}^{\mathrm{g}}$

Note that the solution of the final step provides a feasible solution even though the resulting model in each step is a linear program. This is because all integer variables are set to integer values in step 4 despite being defined as continuous variables.

Table S15 shows the performance of FSA against the exact approach (EA), which is the original MILP model solved by Gurobi. We run each instance for 12 hours. All gaps are calculated based on the *best lower bound obtained* by Gurobi within the time limit. This lower bound is usually the solution for the LP relaxation of the problem.

Across both decarbonization targets, only three instances with two and three representative days are solved with less than a 1% gap within the time limit. EA is unable to obtain any incumbent solution for instances with seven, nine and 10 representative days under 80%, and instances with eighth, 11 and 12 representative days under 95% decarbonization target. EA can only achieve better gaps in instances with less than five days. On average, the FSA achieves a gap of about 4.6% and 5.6% for 80% and 95% decarbonization goals whereas these gaps stand at 57% and 42% for EA. Notice that the actual gap between the feasible solution obtained by FSA and the *optimal* solution is probably much lower.

Table S15: Performance of the FSA vs. the exact approach (EA) for HX electrification scenario with exemplar weather year of 2003. Gaps are calculated based on the best lower bound found by Gurobi within the time limit.

	Decarbonization Target=80%				Decarbonization Target=95%			
	EA		FSA		EA		FSA	
$\lvert\mathfrak{R}\rvert$	CPU (h)	Gap (%)	CPU (h)	Gap (%)	CPU (h)	Gap (%)	CPU (h)	Gap (%)
2	4.4	0.2	0.1	3.1	1.6	0.5	0.1	2.9
3	12	0.5	0.1	6.1	12	3	0.1	7.1
4	12	92	0.1	9.8	12	74	0.1	4.7
5	12	41	0.2	3.9	12	2	0.1	2.2
6	12	34	0.3	2.8	12	9	0.2	3.8
7	12	-	0.4	2.7	12	73	0.3	5.6
8	12	93	0.4	5.2	12	-	0.5	13.3
9	12	-	0.5	3.9	12	86	0.4	5.7
10	12	-	0.6	3.1	12	92	0.5	5.8
11	12	93	0.7	5.6	12	-	1.3	5.8
12	12	99	0.5	4.6	12	-	0.9	5.1
ave.	11.3	57	0.4	**4.6**	11.1	42	0.4	**5.6**

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

12 Bottom-Up Building Energy Model

12.1 ResStock Modeling

To generate the most granular level of load data in this study, we employ NREL's ResStock tool [25]. We use ResStock to produce two key outputs: heating-and-cooling-related thermal loads, which are the thermal energy required to maintain comfort in the living space, and non-heating-and-cooling-related fuel demands such as gas demand for water heating. At the time of conducting the analysis for this paper, the most recent release version of ResStock (2.5) did not support hourly heating-and-cooling-related thermal loads as a model output. Therefore, we used a clone of the develop branch on ResStock's GitHub, which with more recent updates included these features.

Our modeling relies on archetypes that collectively represent the residential building stock. Because each archetype represents a group of many identical homes that may receive varying upgrades, we define so-called *sub-archetypes* to represent each of these variants. The *baseline sub-archetype* is the original version of the archetype, with no upgrades. The *electrified sub-archetypes* capture different possible mixes of heat pump sizing and envelope upgrades applied to the baseline sub-archetype (which we refer to as *upgrade packages*). The electrified sub-archetypes are as follows:

1. Summer-sized heat pump, no envelope upgrades
2. Summer-sized heat pump, with envelope upgrades
3. Winter-sized ("whole-home") heat pump, no envelope upgrades
4. Winter-sized ("whole-home") heat pump, with envelope upgrades

It should be noted that summer-sized heat pumps use the existing system as a backup heating source during the coldest periods. This is typically a hybrid heating system because most of the stock has existing fuel-based heating.

12.1.1 Building Stock Turnover Modeling

The input parameters of ResStock include housing stock data that is presented in the form of probability distributions for a range of various interdependent building characteristics. In order to reflect the 2050 housing stock, we alter the default ResStock distributions to approximate the projected housing stock for 2050. We generalize the projected housing stock changes in Massachusetts, described in "Building Sector Report of the Massachusetts 2050 Decarbonization Roadmap Study" [113], to the entirety of New England. When normalized for floor area, the Roadmap projects that the cohort of residential buildings constructed after the present day will be 23% of the stock in 2050. We modify the ResStock probability distributions to reflect this growth by proportionally increasing the likelihood that the ResStock model samples an archetype of 2010s vintage, the most recent vintage available in ResStock. This implicitly increases the proportions of homes in our 2050 baseline stock that have newer construction characteristics, such as high-quality insulation. The proportions of older vintages in the remainder of the stock are assumed unchanged. In addition to overall stock turnover, we also incorporate projected changes in building type – for example, the Roadmap projects increasing rates of construction for large multi-family buildings in the coming decades. However, we do not include representation of routine building upgrades (e.g. appliance replacement, window replacement) that may occur in existing homes between the present day and 2050.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Table S16: Comparison of upgrades specified in the Massachusetts Decarbonization Roadmap versus the equivalents that we modeled in ResStock.

Upgrade	MA Decarbonization Roadmap	ResStock equivalent
Roof/Ceiling insulation	R-60	R-60
Wall (fill) insulation	R-15	R-13
Sheathing insulation	N/A	R-5
Rim joist insulation	N/A	R-13
Foundation wall insulation	N/A	R-10
Infiltration reduction	0.4 CFM/sf at 0.3 in. wc.	2.25 ACH50

12.1.2 Generating thermal load data

Thermal loads for heating and cooling in the baseline sub-archetypes are necessary inputs for our modeling of the electrified sub-archetypes. ResStock provides data for heating and cooling thermal loads if and only if heating and cooling systems are present, which is not the case for every baseline sub-archetype in New England. To retrieve these loads, we use ResStock to apply an upgrade to every baseline sub-archetype such that it has a heating and cooling system. In particular, we apply an upgrade of a typical single-speed air-source heat pump (in ResStock: "HVAC Heating Efficiency|ASHP, SEER 15, 9.0 HSPF"). ResStock properly sizes the systems such that they meet thermal comfort requirements. Rather than outputting thermal loads, ResStock outputs the heating and cooling energy supplied by the system (*delivered loads*). We accordingly assume that the delivered loads are roughly equivalent to the thermal loads. ResStock also provides the fuel demands corresponding to the new air-source heat pump upgrade, however, we opt to use a different method as the ResStock method largely relies on inefficient electric resistance backup heating. Instead, we develop a heat pump model that more accurately approximates industry standard sizing methods for cold-climate regions such as New England (details in Section 12.2).

12.1.3 Envelope improvements

Our scenarios consider envelope improvements that refer to all post-construction upgrades made to the building exterior to reduce heat loss and improve thermal efficiency. Examples of these activities include adding insulation and sealing air leaks. We use ResStock to generate thermal loads in the presence of basic envelope upgrades for both the baseline and electrified sub-archetypes. We approximately align these improvements with the improvements specified in the "ECM2 – Medium Efficiency" package outlined in the Building Sector Report of the Massachusetts 2050 Decarbonization Roadmap Study [113], which includes the lowest degree of envelope improvements considered in the report, and is thus what we assume to be a basic level of envelope retrofit. ResStock offers a discrete list of improvement options that do not necessarily match a level shown in the Roadmap. We show the comparisons in Table S16 .

The Roadmap only lists a generic "wall insulation" characteristic of R-15, while ResStock takes separate insulation values for wall fill, sheathing, rim joist, and foundation wall insulation. Hence, we select ResStock insulation values such that the combination of these components across the wall section will equal roughly R-15. A given improvement is only applied to a home if it is related to a characteristic the

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

home actually has (e.g., a basement) and the existing option is less efficient. The Roadmap's ECM2 package also includes Energy Recovery Ventilation (ERV), which we do not model as it was not a capability of ResStock at the time we conducted our analysis.

12.1.4 Water heating improvements

In addition to an air-source heat pump, we assume that electrified sub-archetypes also receive a heat-pump water heater (HPWH). We model this in ResStock by selecting a 66-gallon HPWH with a Uniform Energy Factor (UEF) of 3.35. Beyond space and water heating, we do not consider the electrification of other residential end uses. However, space and water heating make up the vast majority of residential electrifiable energy demand [114].

12.2 Heat pump model

We use the thermal loads provided in ResStock as inputs to our modeling of heating and cooling demand for the electrified sub-archetypes. ResStock provides thermal loads for a single living space representing the entirety of the home. Therefore, our models assume the home will be heated by a single heat pump, although in reality some households may opt for multiple smaller heat pumps.

12.2.1 Data underlying heat pump model

Our ASHP model is based on applying statistical linear regression on the large "Cold-Climate Air Source Heat Pump" dataset from NEEP [115]. The NEEP dataset includes data for thousands of cold-climate heat pumps that are submitted by manufacturers and certified by AHRI. Key characteristics of heat pump performance and the associated fuel demands rely on the operating conditions. In our models, we primarily consider the effects of outdoor air temperature (i.e., ambient temperature). This is common in the literature [11, 12]. The NEEP dataset provides capacity and efficiency values at multiple temperatures for all ASHPs listed in the dataset. We preprocess the dataset to form the basis of our regression model. First, we filter for models with Heating Seasonal Performance Factor (HSPF) equal to 10, the level currently required to qualify for electrification incentives in Massachusetts [116]. In order to avoid extrapolating beyond the range of our regression data, we then filter for models for which the manufacturer provided data for performance below -15 ^{o}C (5 ^{o}F). We then remove duplicate data which leads to some heat pumps being over-represented in the dataset. This results in our final model being based on a dataset of 161 different heat pump models. We note that because some heat pumps in the dataset are likely to be more popular than others, this does not result in an aggregate of the most likely heat pump to be adopted, but rather an approximation of the "average" HSPF 10 heat pump on the market.

12.2.2 Coefficient of Performance (COP) Model

The hourly COP of the heat pump determines the ratio of the useful thermal energy supplied to (or removed from) the space to the electricity consumption in a given hour.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

$$\mathrm{COP}_{\mathrm{heat}} = \frac{|\dot{Q}_{\mathrm{out}}|}{\dot{E}_{\mathrm{in}}} \tag{22}$$

The coefficient of performance (COP) for heat pumps in heating mode decreases in colder temperatures. Similarly, warmer temperatures adversely impact the COP of heat pumps in cooling mode. We compute the hourly COPs as a function of the hour's outdoor air temperature. In addition to temperature, COP varies with the "part-load ratio" which is the amount of energy the heat pump supplies relative to its maximum capacity at the operating outdoor air temperature. To simplify the model, we do not consider part-load performance, an assumption made in similar studies [11]. The temperature-vs.-COP function is based on a least-squares linear regression of the COP and temperature values listed in the NEEP dataset. The NEEP dataset lists multiple COPs for a given temperature depending on the part-load ratio of the heat pump. We use the COP values for heat pump operation at maximum capacity. There are separate COP curves for heating and cooling:

$$\mathrm{COP}_{\mathrm{heat,h}} = f(T_h) = 0.045T_h + 2.73 \tag{23}$$

$$\mathrm{COP}_{\mathrm{cooling,h}} = f(T_h) = -0.116T_h + 7.35 \tag{24}$$

where T_h is the parameter for the hourly temperature in degrees Celsius. Our linear model for heating mode, as shown in Fig. S25, is inherently a simplification of reality. According to performance testing, the dependence of COP on temperature is non-linear [117]. For our heating COP dataset, linear and quadratic fits result in essentially equivalent curves, suggesting that additional model complexity would not necessarily result in better fit at the expense of tractability. Similarly, Fig. S26 shows the regression model for COP in cooling mode.

12.2.3 Heating Capacity Derating Model

The *capacity* of a heat pump refers to the maximum heating or cooling output it can produce. Similar to COP, the capacity depends on temperature such that colder temperatures generally reduce the heating capacity of a heat pump relative to warmer temperatures, and vice versa. Colder temperatures typically imply higher heating loads within a building; because this coincides with reduced heating capacity in the heat pump, there is a temperature range for which heating supply cannot match heating load. The modeling of capacity derating enables us to identify hours in which the heat pump's capacity is less than the heating load, hence backup heating may be necessary (discussed in more detail in Section 12.2.5). Therefore, our model incorporates temperature-related capacity derating for heating. In addition to COP values at each temperature, the NEEP dataset contains heating capacity values. For the ASHP models represented in the dataset, there are *maximum* and *minimum* capacity values provided for each temperature. Such a range of capacities can be present in a variable speed system, which has a compressor speed that can be modulated via controls, meaning there can be a range of capacities for a given temperature. We assume that for a given heat pump model represented in the dataset, the capacity value at a given temperature is the maximum value listed in the dataset. To obtain a model for the capacity derating, we take a regression of the capacity values across the dataset for all models, normalized compared to the capacity at 8.3 ^{o}C (47 ^{o}F) of each respective model. This enables us to obtain a slope that represents the average percentage loss of heating capacity for every degree Celsius drop in temperature, for

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

all heat pumps represented in the dataset. The capacity value for a given hour, C_h, is calculated as

$$C_h(T_h) = (1 - 0.0153 \cdot (T_{\text{sizing}} - T_h)) \cdot C_{\text{sizing}} \tag{25}$$

We first define a *sized capacity*, C_{sizing}, the capacity at which the heat pump is sized at the *sizing temperature*, T_{sizing}. The method for sizing is discussed in Section 12.2.4, and differs depending on the sizing method. We assume the decrease in the capacity below T_{sizing} is proportional to the slope obtained from the regression in our capacity derating model. Note that the logic also applies in reverse, e.g., if we assume that the heat pump is running at an outdoor air temperature 10 degrees above the sizing temperature, the heat pump has a capacity 15.3% higher than the capacity at the sizing temperature. Similar to the COP model, our linear model for heating capacity abstracts away from potential non-linearities in the temperature dependence of heating capacity, which is empirically the case [117]. Our model is depicted visually in Fig. S27.

As a way to simplify our model, we do not consider capacity derating for cooling. The NEEP capacity data suggests that cooling capacity declines only modestly with increases in temperature. Additionally, given the dominance of heating loads over cooling loads in our study area, cooling capacity derating is less relevant to determining important hourly demand phenomena such as peak electricity loads.

12.2.4 Sizing Model

A key consideration in installing a heat pump is its *size*, that is, the capacity. In our analysis, we define two possible sizing methods during an installation: sizing a smaller heat pump, primarily for cooling (summer sizing) and sizing a larger heat pump for heating the whole home (winter sizing). Both systems provide some amount of heating in the winter, with the winter-sized system intended to heat throughout the year. For each archetype, we size heat pumps for summer and winter approximately according to current and proposed *ACCA S* methods, a set of industry-standard guidelines for HVAC sizing [118]. As inputs to our sizing methods, we leverage data for the typical meteorological year (TMY) [119] to produce archetype loads and temperatures from ResStock corresponding to typical weather conditions. We do this as part of an attempt to approximate the ACCA S sizing methods' usage of long-run weather averages to determine the design conditions. Our sizing methods differ from ACCA S in that we use the delivered heating or cooling loads (see Section 12.1.2) rather than so-called "design loads," both of which are modeled outputs, but generated through different methods. We do this because ResStock did not provide design loads for winter-sized heat pumps as an output at the time of our analysis.

Winter sizing

In the case of winter sizing, the size of the installed heat pump is primarily determined by the heating load. For each archetype, we first take the 99th percentile hourly heating load from the year of TMY data and "size" the heat pump such that it can provide this capacity C_{sizing} at the sizing temperature, T_{sizing}. In the case of winter sizing, we use the 1st percentile temperature. Because the 99th percentile load often occurs at temperatures above the 1st percentile temperature, many of our archetypes effectively meet 100% of the load across the typical meteorological year. The cooling capacity is tied to the heating capacity. Generally, the nameplate heating and nameplate cooling capacity of a heat pump are similar, where the nameplate

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

heating capacity is the capacity at 8.3 ^{o}C (47 ^{o}F) and the nameplate cooling capacity is the capacity at 35 ^{o}C (95 ^{o}F). Analysis of our dataset shows that these values are on average within about 4% of one another. Therefore, we assume the heating and cooling nameplate capacities are equal. Because we do not model cooling capacity derating, we effectively set the temperature-invariant cooling capacity as equal to the derated capacity of the heating system at 8.3 ^{o}C (47 ^{o}F), that is, $C(8.3^o\text{C})$. Because we only use a single capacity curve, our heat pump models and the associated sizing methods do not consider the potential issue of "short-cycling," which is rapid on-off switching that may occur when the heating or cooling demand is below the heat pump's minimum capacity at the operating temperature. Although variable speed systems offer a range of capacities at any given temperature, short cycling may still occur, particularly in winter-sized systems, leading to humidity control issues that would generally be desired to be avoided in the sizing calculations. We assume the combination of our sizing methods and heat pump operation does not result in excessive short-cycling that would necessitate different installation configurations.

Summer sizing

In the case of summer sizing, the heating capacity of the installed heat pump is tied to the summertime cooling load (kW). For each archetype, we calculate the 99th percentile cooling load using the TMY weather data. We then size the cooling capacity of the heat pump to provide 1.3x this value (kW) at T_{sizing} where T_{sizing} is the 99th percentile temperature. We then set the nameplate heating capacity equal to the cooling capacity. The 130% factor approximates the ACCA S method for sizing a heat pump primarily used for cooling in a heating-dominated climate. Hence, we use it as the basis of our method for "summer-sized" systems.

12.2.5 Backup heating for "summer-sized" systems

In our primary modeling scenarios, we consider the usage of backup heating systems, particularly in the case of summer-sized systems that are not sized to meet heating loads in the winter. The possible configurations of a backup heating system are typically constrained by the nature of the existing heating system and the type and size of heat pumps installed. We refer to summer-sized systems for which the backup system is fueled by non-electric fuels as a "hybrid" system. We define our configurations based on those used in NREL's End-Use Savings Shapes (EUSS) project [120].

Configuration 1: Ducted home w/ existing backup

For electrified sub-archetypes where the corresponding baseline sub-archetype has an existing ducted system, we assume the home will have installed a ducted heat pump downstream of the existing furnace or boiler that will serve as backup. The existing system may be fueled by gas, fuel oil, electricity, or such other fuels as wood or propane. In this case, the heat pump and the existing system cannot run at the same time. Installers often define a switchover temperature, $T_{switchover}$ below which the heat pump becomes relatively inefficient and has reduced capacity, where the heat pump is deactivated and the backup system meets the entire load. We define $T_{switchover}$ as 41^oF (5^oC) in summer-sized systems, in line with NREL's assumption for an existing backup system in the EUSS study. The efficiency of the backup system is treated as temperature-invariant and is taken from the features of the corresponding baseline sub-archetype, which include the nominal efficiency of the existing system. Although true for only a small portion of homes in New

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

England, the existing system may be electric resistance.

Configuration 2: Ductless home w/ existing backup

For archetypes where the existing heating system is ductless, such as those with water-based heating distribution systems, the heat pump and existing system generally will not interfere with one another and can run simultaneously. In this case, we assume sensors and controls have been installed such that, when the heat pump is unable to meet the heating load, the backup system runs to make up the difference between the heating load and the heat pump's capacity. This is likely to be necessary for summer-sized systems during the winter.

Configuration 2a: Ductless or Ducted home w/ electric backup

While our scenarios in the main text assume all summer-sized heat pumps are configured with the existing sustem as backup, our model also includes logic to reflect newly-installed electric resistance backup heating, as evaluated in section 7.2. In this case, we assume the electric resistance coils have been installed downstream of the heat pump in a ducted system or separate from the existing system in a ductless system (e.g., by installing electric baseboard heating). The heat pump and backup system can run simultaneously as in Configuration 2. We treat the electric resistance backup as having a temperature-invariant efficiency of 100%.

12.3 Aggregation modeling

12.3.1 Aggregation method

For each archetype, a weight W is applied, equal to the number of homes the archetype represents in the residential stock. Additionally, each archetype in a given zone z has the same weight. We determine the overall number of homes represented by a given archetype by dividing the projected number of homes H_z in the zone by the number of archetypes in the zone, A_z:

$$W_z = H_z / A_z \tag{26}$$

A_z is equal to roughly 400 for every zone in our analysis.

For each archetype in this analysis, there is a single baseline sub-archetype and four electrified sub-archetypes considered, each with its own weight. Let W_u correspond to the weight of a sub-archetype with a given upgrade package in the set of possible upgrade packages U. One such u is the baseline home b, represented by weight W_b. The electrified sub-archetypes correspond to the remaining weights W_u:

1. W_s: Summer-sized heat pump, no envelope upgrades
2. W_{se}: Summer-sized heat pump, with envelope upgrades
3. W_w: Winter-sized heat pump, no envelope upgrades
4. W_{we}: Winter-sized heat pump, with envelope upgrades

Let a given archetype in zone z be archetype i, which is in the set of all archetypes in zone z, I_z. For a given archetype i, the sum of the weights of all its sub-archetypes is equal to the weight of the archetype, which is equal to the shared weight of all the archetypes in the zone:

$$W_i = \sum_u W_{i,u} = W_{i,b} + W_{i,s} + W_{i,se} + W_{i,w} + W_{i,we} = W_z \quad \forall\, i \in I_z \tag{27}$$

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

The weights for the sub-archetypes can be altered for any given archetype to reflect the level of heat pump deployment and mixing of sizing methods in the portion of the stock represented by the archetype. The profiles for each sub-archetype consist of the heating-and-cooling-related and all other demands. The hourly profiles for each fuel type f for zone z are equal to the summation of the profiles for all sub-archetypes in the zone multiplied by their respective weights. If $L_{f,i,u,h}$ is the demand for fuel f for archetype i with upgrade package u in hour h, then the hourly demand for fuel f in zone z is:

$$\sum_{i=1}^{n} W_{i,b}L_{f,i,b,h}+W_{i,s}L_{f,i,s,h}+W_{i,se}L_{f,i,se,h}+W_{i,w}L_{f,i,w,h}+W_{i,we}L_{f,i,we,h} = L_{f,h} \tag{28}$$

12.3.2 Heat pump deployment scenarios

We base our heat pump deployment scenarios on the Massachusetts Clean Energy and Climate Plan (CECP) for 2050, with our Medium Electrification scenarios corresponding to the CECP "Hybrid" scenario and the High Electrification scenarios corresponding to the CECP "High Electrification" scenario [7]. The CECP adoption projections include separate adoption rates for systems of different sizes and are listed in units of heating systems. On advice from the Massachusetts Executive Office of Energy & Environmental Affairs, we assume the number of homes that have adopted heat pumps is equal to the number of heat pumps in the overall heating stock. The CECP scenarios do not include projections for adoptions of envelope improvement for 2050; hence, for our envelope upgrade scenarios, we assume that 70% of all electrified homes receive envelope upgrades. The deployment scenarios manifest themselves as the weights. In our modeling, we assume all regions have levels of adoption in line with the CECP scenarios. We also assume the heat pump deployment scenarios are such that every archetype receives the same proportions of upgrade packages (i.e., proportions of electrified and baseline sub-archetypes), with the exception of archetypes for which the baseline sub-archetype already has a heat pump, which do not receive upgrades at all. For example, if we assume 75% of homes in Zone z get new winter-sized heat pumps with envelope improvements:

$$W_{i,we} = 0.75W_i = 0.75W_z \tag{29}$$

If Wz is 400 homes, then 300 of the homes represented by archetype i are represented by sub-archetype i, we. If there are 50 archetypes total for Zone z that do not already have a heat pump in their baseline sub-archetype, then 300 homes represented by each archetype receive upgrade package we and in total 15,000 homes in the zone receive upgrade package we. Our reference scenario is taken from NREL's Electrification Futures study's "High Electrification – Moderate Technology Advancement" scenario [9]. Similar to the CECP, it presents projections in terms of heating system stock numbers; however, it does not include information on sizing. As our most conservative case, we assume heat pumps deployed in this scenario are summer-sized (including existing backup) and that no envelope upgrades are applied. Table S17 shows the scenarios in a table.

For a small number of archetypes in our projected 2050 stock, the baseline sub-archetype already has a heat pump modeled in ResStock, which by default fits under Configuration 2a (see Section 12.2.5). We assume these homes contribute to the

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Table S17: Electrification scenarios in tabular form. Units in percent of total homes in the region.

	RF	ME	MX	HE	HX
Summer-sized	6.2	41.2	12.4	16.8	5.0
Summer-sized + Eff Improv	0	0	28.8	0	11.8
Winter-sized	0	16.1	4.8	62.1	18.6
Winter-sized + Eff Improv	0	0	11.3	0	43.4

projections of adoption. We lump these into the "summer-sized" category in Fig. 1 in the main text, although they would exhibit electricity demands in excess of the "winter-sized" systems due to high reliance on electric resistance heating.

12.3.3 Population and household data

Our analysis calls for projections of the number of homes in 2050. Generally, state-level projections for home growth through 2050 are unavailable. We assume the number of homes grows proportionally to households or population, depending on data availability. Massachusetts provides household count projections as part of its Decarbonization Roadmap [113]. For other states, we generally source state-level population growth projections from state agencies [121–124]. We also assume that each county's number of homes will grow at a rate equal to the state-wide projected growth. In some states, projections only extend to 2040. In these cases, we linearly extrapolate the growth through 2050. Depending on the zone, each archetype represents between 600 and 2,000 homes.

12.3.4 Modeling of present-day demand

In addition to the reference case, it is useful to have a current baseline against which to compare the hourly demands of the future scenarios. We refer to this case as the "present-day" case. Although historical aggregate demands such as annual demands are available for the residential sector [3], residential hourly demands generally are not. Hence, we attempt to emulate these demands using the same workflow as the electrification scenarios. Similar to the future scenarios, we simulate 400 archetypes per zone, however, we use ResStock's default housing dataset representative of the 2018 stock. We use weather data for 2001-2020 without any climate change adjustments applied. We determine our archetype weights using home count data for 2020. In this data, there is no modeling of electrification impacts. Similar to the demand modeling for 2050, where the 2001-2020 weather data is used to represent individual possible weather pattern realizations that are then adjusted for climate change, the present-day demand data can be seen as representative of demands for the 2020 baseline under a portfolio of 2001-2020 weather patterns. Implicitly, these results assume there are no climate change effects that would significantly impact weather patterns between 2001-2020.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

12.4 Demand-side weather and climate modeling

12.4.1 Weather data

Hourly weather data is a key input to the bottom-up method. We collect hourly actual meteorological year (AMY) data for 2001-2020 in 44 locations across New England corresponding to the weather stations in the DOE's TMY3 dataset, shown in Fig. S28. Although the model can support higher levels of spatial granularity, in this analysis we are limited to using weather data from the locations within the TMY3 dataset because our heat pump sizing calculations require TMY data as an input before we can run any building energy simulations using AMY data. Each archetype is assigned the weather data closest to the county in which the archetype lies.

We source the TMY data used in this analysis from the DOE TMY3 dataset [54]. Our AMY data was provided by OikoLab, which furnish an API for more efficiently accessing large amounts of weather reanalysis data produced by the ERA5 project [48]. The ERA5 data is also openly available on the ERA project website [47]. We elaborate more on our rationale for selecting the ERA5 data below.

Selecting the ERA5 dataset

Our weather data needs for the bottom-up analysis are particularly constrained. We are looking for multi-decadal dataset at an hourly resolution, that also contained the several hourly weather parameters needed for building energy demands via the ResStock tool. In addition, we sought a dataset with high spatial granularity, which would be useful for capturing weather heterogeneity in a relatively small region like New England.

We generally lack the expertise to assess the comparative merit or accuracy of different historical weather datasets based on the parameters alone, an area for which there is a rapidly growing literature among climate scientists [125–127]. However, because outdoor air temperature is the most impactful weather-related determinant of demand, we compared the temperature profiles among two datasets that we identified as meeting our constraints, the ERA5 [47] and MERRA2 [128] datasets. In Fig. S29, we show comparisons of the two datasets as benchmarked against observation data provided by NOAA [129] for a handful of locations in different subregions of New England (namely: Boston, MA; Burlington, VT; and Bridgeport, CeiT). We also compare the datasets to the weather data used by NREL in the End Use Load Profiles project [25]. To capture the differences across different time periods, we compare the accuracy of the average daily temperature profiles for the winter months (January, February, November, and December) and summer months (May, June, July, and August) in both 2012 and 2018.

Immediately noticeable is that the NREL weather data fits the NOAA data the best. This is because this dataset is constructed primarily from the NOAA data, with some interpolation applied for missing observations [25]. However, the NREL data could not be a candidate dataset for our analysis because NREL has only published the data with limited features and for just two weather years. Among the ERA5 and MERRA2 datasets and across the locations, it can be seen that the ERA5 temperature profile fits significantly better, regardless of location or year, suggesting that ERA5 is the better data source for this particular study area. For the locations and seasons shown, the ERA5 data exhibits an average absolute temperature error of 0.4 to 1.3 C, whereas the MERRA2 data exhibits an error of 1.7 to 3.5 C.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

The full set of temperature distributions for Boston across the 20 weather years used in our electrification scenarios (based on the morphed ERA5 data) are shown in Fig. S30 below.

12.5 Morphing method

The 2050 time horizon of our analysis necessitates consideration of climate change effects on demand patterns. Increases in temperatures and changes in other meteorological conditions will likely affect demand [130, 131]. Building energy models like the EnergyPlus model underlying ResStock require several specific weather variables at hourly resolution. Climate models generally do not produce predictions at hourly resolution for these variables, an issue that is broadly recognized in the literature [132], instead providing predictions at the daily or monthly resolution. The "morphing method," first introduced in [49], is a particularly common method for overcoming this disconnect in the literature, combining hourly historical weather data with monthly climate models to produce realistic weather patterns that reflect the long-run effects of climate change [133, 134]. The morphing method offsets and scales the hourly values within each month to reflect the monthly average changes projected by the climate model. It also manipulates the hourly values to reflect changes in the monthly maxima and minima when they are available. We use the morphing method to apply the effects of climate change to our baseline 2001-2020 weather data. As described in [49], the selection of the morphing operation depends on the nature and units of the underlying variable. For example, it is appropriate to "shift" mean temperatures because they are in absolute units of C, but relative humidity, which is provided in percentages and cannot go below zero, is more amenable to a "scale" operation. We select the CESM2 model [52] as the basis for the morphing because of its popularity, because its data is accessible through the CMIP6 project [135], and because it provides the necessary variables as outputs. We select the SSP3-70 emissions scenario as an approximation of medium-to-high warming. In determining the long-run changes from the climate model, we compare the average outputs of the model between the periods 2015-2023 and 2046-2054. As opposed to simply comparing the outputs for 2020 to outputs for 2050, taking averages of multi-year periods smooths out any changes that may be the result of the interannual weather variation simulated in CESM2 rather than long-run climatic changes. We note that because our baseline weather data forced through the morphing method is taken from any year 2001-2020, we inherently assume no differences among the 2001-2020 weather years due to the effects of climate change.

A variable required by EnergyPlus but not provided in the CESM2 model is the dew point temperature. We approximate it from the morphed dry bulb temperature and relative humidity values using the MetPy package in Python [136].

References

[1] U.S. Environmental Protection Agency (EPA). Sources of Greenhouse Gas Emissions. website (2023). URL https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions. Accessed: 2023-3-2.

[2] U.S. Energy Information Administration (EIA). U.S. CO2 emissions from energy consumption by source and sector, 2021.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

Table S18: EnergyPlus/ResStock variable to CMIP6 standard variable mappings

EnergyPlus/ResStock variable	**CMIP6 Standard variable name(s) used in morphing**	**Monthly morphing operation applied**
Dry bulb temperature	tas, tasmin, tasmax	"Shift" mean temperatures, "stretch" distance between max & min temperatures [49]
Wind speed	sfcWind	"Stretch" as in [50]
Relative humidity	hurs	"Stretch" as in [50] while keeping bounded between 0-100%
Atmospheric pressure	ps	"Shift" as in [49]
Total sky cover	clt	"Shift" as in [49], while keeping bounded between 0-100%
Global horizontal radiation, diffuse horizontal radiation, direct normal radiation horizontal IR radiation	rsds	"Stretch" based on global horizontal radiation as in [51] due to projections only existing for this variable

website (2021). URL https://www.eia.gov/energyexplained/energy-and-the-environment/images/CO2-spaghettichart-2021.pdf. Accessed: 2022-5-15.

[3] U.S. Energy Information Administration (EIA). Natural Gas Residential Consumption, Annual. website (2023). URL https://www.eia.gov/opendata/browser/co2-emissions. Accessed: 2023-3-2.

[4] International Energy Agency (IEA). Net Zero by 2050 A Roadmap for the Global Energy Sector. website (2022). URL https://www.iea.org/reports/net-zero-by-2050. Accessed: 2023-3-2.

[5] Gaur, A. S., Fitiwi, D. Z. & Curtis, J. Heat pumps and our low-carbon future: A comprehensive review. *Energy Research & Social Science* **71**, 101764 (2021) .

[6] Internal Revenue Service. Frequently asked questions about energy efficient home improvements and residential clean energy property credits (2022). URL https://www.irs.gov/pub/taxpros/fs-2022-40.pdf. Accessed: 2023-04-18.

[7] Massachusetts Department of Energy Resources. Massachusetts Clean Energy and Climate Plan for 2050. website (2022). URL https://www.mass.gov/doc/2050-clean-energy-and-climate-plan/download. Accessed: 2023-3-2.

[8] American Council for an Energy-Efficient Economy (ACEEE). State Policies and Rules to Enable Beneficial Electrification in Buildings

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

through Fuel Switching. website (2022). URL https://www.aceee.org/sites/default/files/pdfs/state_fuel-switching_policies_and_rules_7-21-22.pdf. Accessed: 2023-3-2.

[9] Mai, T. *et al.* Electric Technology Adoption and Energy Consumption (2018). URL https://www.osti.gov/servlets/purl/1461472/.

[10] Deetjen, T. A., Walsh, L. & Vaishnav, P. US residential heat pumps: the private economic potential and its emissions, health, and grid impacts. *Environmental Research Letters* **16** (8), 084024 (2021) .

[11] Vaishnav, P. & Fatimah, A. M. The environmental consequences of electrifying space heating. *Environmental Science & Technology* **54** (16), 9814–9823 (2020) .

[12] Waite, M. & Modi, V. Electricity load implications of space heating decarbonization pathways. *Joule* **4** (2), 376–394 (2020) .

[13] Wu, Y., Waite, M., Conlon, T. & Modi, V. The role of electrification induced peak loads and gas infrastructure constraints on decarbonization pathways in new york state. *Energy Strategy Reviews* **44**, 100985 (2022) .

[14] Khorramfar, R., Mallapragada, D. & Amin, S. Electric-gas infrastructure planning for deep decarbonization of energy systems. *Applied Energy* **354**, 122176 (2024) .

[15] Von Wald, G., Sundar, K., Sherwin, E., Zlotnik, A. & Brandt, A. Optimal Gas-Electric Energy System Decarbonization Planning. *Advances in Applied Energy* 100086 (2022) .

[16] Deakin, M. *et al.* Impacts of heat decarbonization on system adequacy considering increased meteorological sensitivity. *Applied Energy* **298**, 117261 (2021) .

[17] Eggimann, S., Usher, W., Eyre, N. & Hall, J. W. How weather affects energy demand variability in the transition towards sustainable heating. *Energy* **195**, 116947 (2020) .

[18] Peacock, M., Fragaki, A. & Matuszewski, B. J. The impact of heat electrification on the seasonal and interannual electricity demand of great britain. *Applied Energy* **337**, 120885 (2023) .

[19] Staffell, I. & Pfenninger, S. The increasing impact of weather on electricity supply and demand. *Energy* **145**, 65–78 (2018) .

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

[20] Perera, A., Nik, V. M., Chen, D., Scartezzini, J.-L. & Hong, T. Quantifying the impacts of climate change and extreme climate events on energy systems. *Nature Energy* **5** (2), 150–159 (2020) .

[21] Pfenninger, S. Dealing with multiple decades of hourly wind and pv time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability. *Applied energy* **197**, 1–13 (2017) .

[22] White, P. R., Rhodes, J. D., Wilson, E. J. & Webber, M. E. Quantifying the impact of residential space heating electrification on the texas electric grid. *Applied Energy* **298**, 117113 (2021) .

[23] Tarroja, B. *et al.* Translating climate change and heating system electrification impacts on building energy use to future greenhouse gas emissions and electric grid capacity requirements in california. *Applied energy* **225**, 522–534 (2018) .

[24] Murphy, C. *et al.* Electrification futures study: scenarios of power system evolution and infrastructure development for the united states (2021).

[25] Wilson, E. *et al.* End-Use Load Profiles for the U.S. Building Stock: Methodology and Results of Model Calibration, Validation, and Uncertainty Quantification (2022). URL https://www.osti.gov/servlets/purl/1854582/.

[26] ISO New England. Energy, Load, and Demand Reports. website. URL https://www.iso-ne.com/isoexpress/web/reports/load-and-demand/-/tree/net-ener-peak-load.

[27] U.S. Census Bureau QuickFacts: Vermont. URL https://www.census.gov/quickfacts/fact/table/VT/PST045222.

[28] U.S. Census Bureau QuickFacts: Suffolk County, Massachusetts. URL https://www.census.gov/quickfacts/fact/table/suffolkcountymassachusetts/PST045222.

[29] Energy and Environmental Economics (E3) Inc. and Energy Futures Initiative (EFI). Net-Zero New England: Ensuring Electric Reliability in a Low-Carbon Future (2020) .

[30] Commonwealth of Massachusetts. MA Decarbonization Roadmap. website (2021). URL https://www.mass.gov/info-details/ma-decarbonization-roadmap. Accessed: 2023-7-16.

[31] MIT Energy Initiative. The Future of Energy Storage (2022).

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

[32] Cole, W. *et al.* Quantifying the challenge of reaching a 100% renewable energy power system for the united states. *Joule* **5** (7), 1732–1748 (2021) .

[33] ISO New England. Resource mix. website. URL https://www.iso-ne.com/about/key-stats/resource-mix.

[34] Brown, P. R. & Botterud, A. The value of inter-regional coordination and transmission in decarbonizing the US electricity system. *Joule* **5** (1), 115–134 (2021) .

[35] Dowling, J. A. *et al.* Role of long-duration energy storage in variable renewable electricity systems. *Joule* **4** (9), 1907–1928 (2020). URL https://www.sciencedirect.com/science/article/pii/S2542435120303251. https://doi.org/https://doi.org/10.1016/j.joule.2020.07.007 .

[36] Ruhnau, O. & Qvist, S. Storage requirements in a 100% renewable electricity system: Extreme events and inter-annual variability. *Environmental Research Letters* **17** (4), 044018 (2022) .

[37] Shaner, M. R., Davis, S. J., Lewis, N. S. & Caldeira, K. Geophysical constraints on the reliability of solar and wind power in the united states. *Energy & Environmental Science* **11** (4), 914–925 (2018) .

[38] Joint LDCs. Massachusetts DPU 20-80 - The LDC's Common Regulatory Framework. website (2022). URL https://eeaonline.eea.state.ma.us/DPU/Fileroom/dockets/bynumber/20-80. Accessed: 2023-5-20.

[39] American Gas Association. Implications of Policy-Driven Residential Electrification (2018). URL https://www.aga.org/research-policy/resource-library/implications-of-policy-driven-electrification/.

[40] Strbac, G. *et al.* Analysis of alternative uk heat decarbonisation pathways. *Imperial College London: London, UK* (2018) .

[41] Tonn, B., Rose, E. & Hawkins, B. Evaluation of the us department of energy's weatherization assistance program: Impact results. *Energy Policy* **118**, 279–290 (2018) .

[42] Sorrell, S., Dimitropoulos, J. & Sommerville, M. Empirical estimates of the direct rebound effect: A review. *Energy policy* **37** (4), 1356–1371 (2009) .

[43] Sepulveda, N. A., Jenkins, J. D., De Sisternes, F. J. & Lester, R. K. The role of firm low-carbon electricity resources in deep decarbonization of power generation. *Joule* **2** (11), 2403–2420 (2018) .

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

[44] Rosillo-Calle, F. A review of biomass energy–shortcomings and concerns. *Journal of Chemical Technology & Biotechnology* **91** (7), 1933–1945 (2016) .

[45] Grubert, E. At scale, renewable natural gas systems could be climate intensive: the influence of methane feedstock and leakage rates. *Environmental Research Letters* **15** (8) (2020). URL https://dx.doi.org/10.1088/1748-9326/ab9335. https://doi.org/10.1088/1748-9326/ab9335, publisher: IOP Publishing .

[46] (EPA), E. P. A. Methane emissions reduction program. website. URL https://www.epa.gov/inflation-reduction-act/methane-emissions-reduction-program.

[47] ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Tech. Rep.

[48] OikoLab. Weather and climate data for analysts. website. URL oikolab.com.

[49] Belcher, S. E., Hacker, J. N. & Powell, D. S. Constructing design weather data for future climates. *Building services engineering research and technology* **26** (1), 49–61 (2005) .

[50] Dickinson, R. & Brannon, B. Generating future weather files for resilience. *Proceedings of the international conference on passive and low energy architecture, Los Angeles, CA, USA* 11–13 (2016) .

[51] Jentsch, M. F., Bahaj, A. S. & James, P. A. Climate change future proofing of buildings—generation and assessment of building simulation weather files. *Energy and Buildings* **40** (12), 2148–2168 (2008) .

[52] Danabasoglu, G. *et al.* The community earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems* **12** (2), e2019MS001916 (2020) .

[53] Wilson, E. J., Christensen, C. B., Horowitz, S. G., Robertson, J. J. & Maguire, J. B. Energy efficiency potential in the us single-family housing stock (2017).

[54] Bianchi, C. & Fontanini, A. TMY3 Weather Data for ComStock and ResStock (2022). URL https://data.nrel.gov/submissions/156. https://doi.org/10.7799/1756695 .

[55] Brenner, A., Khorramfar, R. & Amin, S. Learning Spatio-Temporal Aggregations for Large-Scale Capacity Expansion Problems. *arXiv preprint arXiv:2303.08996* (2023) .

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

[56] National Renewable Energy Laboratory (NREL). 2021 Annual Technology Baseline (ATB) (2021). URL https://data.openei.org/submissions/4129. Accessed: 2022-2-8.

[57] Sepulveda, N. A., Jenkins, J. D., Edington, A., Mallapragada, D. S. & Lester, R. K. The design space for long-duration energy storage in decarbonized power systems. *Nature Energy* **6** (5), 506–516 (2021) .

[58] U.S. Energy Information Administration (EIA). EIA Website (2022). URL https://www.eia.gov. Accessed: 2022-2-18.

[59] U.S. Energy Information Administration. US States: State Profiles and Energy Estimates (2023). URL https://www.eia.gov/state/.

[60] Staffell, I., Pfenninger, S. & Johnson, N. A global model of hourly space heating and cooling demand at multiple spatial scales. *Nature Energy* 1–17 (2023) .

[61] Weiss, J. & Hagerty, J. M. Achieving 80% GHG Reduction in New England by 2050 (2019).

[62] Alvarez, R. A. *et al.* Assessment of methane emissions from the US oil and gas supply chain. *Science* **361** (6398), 186–188 (2018) .

[63] Mass.gov. Gseps pursuant to 2014 gas leaks act. website. URL https://www.mass.gov/info-details/gseps-pursuant-to-2014-gas-leaks-act.

[64] U.S. Department of Energy Office of Scientific and Technical Information . Greenhouse gases, Regulated Emissions, and Energy use in Technologies Model. website (2021). URL https://www.osti.gov/doecode/biblio/63109. Accessed: 2023-10-28.

[65] Forster, P. *et al.* The earth's energy budget, climate feedbacks, and climate sensitivity (2021) .

[66] Jenkins, J. D. & Sepulveda, N. A. Enhanced decision support for a changing electricity landscape: the GenX configurable electricity resource capacity expansion model (2017) .

[67] Taylor, Z. T., Mendon, V. V. & Fernandez, N. Methodology for Evaluating Cost-Effectiveness of Residential Energy Code Changes (PNNL-21294 8650) (2015). URL https://www.osti.gov/biblio/1773018. https://doi.org/10.2172/1773018 .

[68] Northeast Gas association. Statistical Guide To The Northeast U.S. Natural Gas Industry 2021. website (2021). URL https://www.northeastgas.org/pdf/stat_guide_2021.pdf. Accessed: 2022-3-18.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

[69] Palmintier, B. S. & Webster, M. D. Impact of operational flexibility on electricity generation planning with renewable and carbon targets. *IEEE Transactions on Sustainable Energy* **7** (2), 672–684 (2015) .

[70] Li, C. *et al.* 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 (2022) .

[71] Kotzur, L., Markewitz, P., Robinius, M. & Stolten, D. Time series aggregation for energy system design: Modeling seasonal storage. *Applied Energy* **213**, 123–135 (2018) .

[72] Xu, Y. *et al.* US test system with high spatial and temporal resolution for renewable integration studies 1–5 (2020) .

[73] Dimanchev, E., Hodge, J. & Parsons, J. Two-way trade in green electrons: deep decarbonization of the Northeastern US and the role of Canadian hydropower (2020) .

[74] New England Clean Energy Connect: NECEC. website (2023). URL https:{//}www.necleanenergyconnect.org{\/}. Accessed: 2023-5-2.

[75] Brown, M. *et al.* Regional Energy Deployment System (ReEDS) Model Documentation: Version 2019 (2020).

[76] Breakthrough Energy. U.S. Test System with High Spatial and Temporal Resolution for Renewable Integration Studies. website (2020). URL https://zenodo.org/record/4538590#.YgL7I-rMK5c. Accessed: 2022-2-8.

[77] U.S. Energy Information Administration (EIA). Emission Coefficients. website (2021). URL https://www.eia.gov/environment/emissions/co2_vol_mass.php. Accessed: 2022-2-8.

[78] Raimi, D. Decommissioning US power plants: Decisions, costs, and key issues (2017).

[79] United States Nuclear Regulatory Commission. Backgrounder on Decommissioning Nuclear Power Plants. website (2021). URL https://www.nrc.gov/reading-rm/doc-collections/fact-sheets/decommissioning.html. Accessed: 2022-2-8.

[80] of Ocean Energy Management, B. Lease and grant information. website. URL https://www.boem.gov/renewable-energy/lease-and-grant-information. Accessed: 2023-2-17.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

[81] National Renewable Energy Laboratory (NREL). Electrification Futures Study Load Profiles. website (2022). URL https://data.nrel.gov/submissions/126. Accessed: 2022-8-21.

[82] Cole, W., Frazier, A. W. & Augustine, C. Cost projections for utility-scale battery storage: 2021 update (2021) .

[83] Reimers, A., Cole, W. & Frew, B. The impact of planning reserve margins in long-term planning models of the electricity sector. *Energy policy* **125**, 1–8 (2019) .

[84] (EPRI), E. P. R. I. Optimization of Transmission Line Design Using Life Cycle Costing (2009).

[85] Teletzke, G. *et al.* Evaluation of practicable subsurface CO2 storage capacity and potential CO2 transportation networks, onshore North America 21–26 (2018) .

[86] Blondes, M. S. *et al.* *National assessment of geologic carbon dioxide storage resources: methodology implementation* (US Department of the Interior, US Geological Survey, 2013).

[87] National Academies of Sciences, Engineering and Medicine, Ocean Studies Board *et al.* Negative emissions technologies and reliable sequestration: A research agenda (2019) .

[88] U.S. Energy Information Administration (EIA). EIA Website (2022). URL https://www.eia.gov/naturalgas/data.php. Accessed: 2022-2-18.

[89] Mallapragada, D. S. *et al.* Life cycle greenhouse gas emissions and freshwater consumption of liquefied marcellus shale gas used for international power generation. *Journal of Cleaner Production* **205**, 672–680 (2018) .

[90] Mesko, J. & Ramsey, J. The use of liquefied natural gas for peaking service. 66 (1996) .

[91] ISO New England. U.S. Energy Information Administration (EIA). website (2022). URL https://www.eia.gov/maps/layer_info{-}m.php. Accessed: 2023-1-8.

[92] . Renewable Sources of Natural Gas: supply and emissions reduction assessment. website (2019). URL https://www.gasfoundation.org/wp-content/uploads/2019/12/AGF-2019-RNG-Study-Full-Report-FINAL-12-18-19.pdf. Accessed: 2023-10-26.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

[93] . Potential of Renewable Natural Gas in New York State. website (2022). URL https://www.nyserda.ny.gov/-/media/Project/Nyserda/files/EDPPP/Energy-Prices/Energy-Statistics/RNGPotentialStudyforCAC10421.pdf. Accessed: 2023-10-26.

[94] Independent Consultant Report. The Role of Gas Distribution Companies in Achieving the Commonwealth's Climate Goals. website (2022). URL https://thefutureofgas.com/content/downloads/2022-03-21/3.18.22. Accessed: 2023-10-25.

[95] Hafner, M. & Luciani, G. *The Palgrave Handbook of International Energy Economics* (Springer Nature, 2022).

[96] Chohan, I. M. *et al.* A review on life cycle assessment of different pipeline materials. *Results in Engineering* 101325 (2023) .

[97] Kaiser, M. J. FERC pipeline decommissioning cost in the US Gulf of Mexico, 1995–2015. *Marine Policy* **82**, 167–180 (2017) .

[98] Energy, D. Green river satellite lng facility basis of cost estimation. website (2019). URL https://pscdocs.utah.gov/gas/21docs/2105712/319896DEUExh2.13StltLNGFsblt8-5-2021.pdf. Accessed: 2022-3-18.

[99] U.S. Energy Information Administration (EIA). EIA Website (2022). URL https://www.eia.gov/environment/emissions/co2_vol_mass.php. Accessed: 2022-2-18.

[100] US Census 2019. County Population Totals: 2010-2019. website (2021). URL https://www.census.gov/data/datasets/time-series/demo/popest/2010s-counties-total.html. Accessed: 2021-12-16.

[101] U.S. Energy Information Administration (EIA). Energy-Related CO2 Emission Data Tables. website (2022). URL https://www.eia.gov/environment/emissions/state/. Accessed: 2022-2-8.

[102] Holmgren, W. F., Hansen, C. W. & Mikofski, M. A. pvlib python: A Python package for modeling solar energy systems. *Journal of Open Source Software* **3** (29), 884 (2018) .

[103] Brown, P. R. & O'Sullivan, F. M. Spatial and temporal variation in the value of solar power across united states electricity markets. *Renewable and Sustainable Energy Reviews* **121**, 109594 (2020) .

[104] Hill, D. F. *et al.* Converting snow depth to snow water equivalent using climatological variables. *The Cryosphere* **13** (7), 1767–1784 (2019) .

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

[105] Ryberg, D. & Freeman, J. Integration, Validation, and Application of a PV Snow Coverage Model in SAM (2017).

[106] National Renewable Energy Laboratory (NREL). 2022 Annual Technology Baseline (ATB): Utility-Scale PV (2022). URL https://atb.nrel.gov/electricity/2022/utility-scale_pv.

[107] ISO New England. 2022 ISO New England Variable Energy Resource (VER) Data Series (2000-2021). website (2022). URL https://www.iso-ne.com/system-planning/planning-models-and-data/variable-energy-resource-data/. Accessed: 2023-1-8.

[108] Hsu, S., Meindl, E. A. & Gilhousen, D. B. Determining the power-law wind-profile exponent under near-neutral stability conditions at sea. *Journal of Applied Meteorology and Climatology* **33** (6), 757–765 (1994) .

[109] Musial, W., Beiter, P., Tegen, S. & Smith, A. Potential Offshore Wind Energy Areas in California: An Assessment of Locations, Technology, and Costs (2016). URL https://www.osti.gov/biblio/1338174. https://doi.org/10.2172/1338174 .

[110] Wiser, R. *et al.* Land-based wind market report: 2022 edition (2022).

[111] Khorramfar, R. Github repository (2023). The repository will be public after the review process. The data is available upon request.

[112] Reuther, A. *et al.* Interactive supercomputing on 40,000 cores for machine learning and data analysis (2018).

[113] Buildings Sector Report: A Technical Report of the Massachusetts 2050 Decarbonization Roadmap Study (2020).

[114] Administration, U. E. I. 2015 Residential Energy Consumption Survey. website. URL www.eia.gov/consumption/residential/data/2015/.

[115] Northeast Energy Efficiency Partnerships. ccASHP Specification & Product List. website. URL https://neep.org/heating-electrification/ccashp-specification-product-list. Accessed: 2023-4-18.

[116] Mass Save. Residential Heat Pump Qualified Product List. website. URL https://www.masssave.com/residential/rebates-and-incentives/air-source-heat-pumps/heat-pump-qualified-list. Accessed: 2023-4-18.

[117] Shoukas, G., Bonnema, E., Paranjothi, G., Faramarzi, R. & Klun, L. Performance Assessment of High Efficiency Variable Speed Air-Source Heat Pump in Cold Climate Applications (2022).

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

[118] Davis, W. Correctly Sizing Variable-Capacity Heat Pump Equipment (2022). URL https://hvac-blog.acca.org/correctly-sizing-variable-capacity-heat-pump-equipment/.

[119] Wilcox, S. & Marion, W. Users manual for TMY3 data sets (2008) .

[120] National Renewable Energy Laboratory (NREL). End use savings shapes residential dataset round 1. website. URL https://www.nrel.gov/buildings/end-use-load-profiles.html. Accessed: 2023-4-18.

[121] New Hampshire Department of Business and Economic Affairs. State, county, and municipal population projections: 2020-2050 (2022). URL https://www.nh.gov/osi/data-center/documents/2022-state-county-municipality-projections-final-report.pdf. Accessed: 2023-4-18.

[122] Rhode Island Statewide Planning Program. Rhode island population projections 2010-2040 (2013). URL https://planning.ri.gov/sites/g/files/xkgbur826/files/documents/census/tp162.pdf. Accessed: 2023-4-18.

[123] Connecticut Data Collaborative. Population projections by town. website. URL http://data.ctdata.org/visualization/population-projections-by-town. Accessed: 2023-01-15.

[124] University of Virginia Weldon Cooper Center, D. R. G. National population projections. website (2018). URL https://demographics.coopercenter.org/national-population-projections/.

[125] Tarek, M., Brissette, F. P. & Arsenault, R. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrology and Earth System Sciences* **24** (5), 2527–2544 (2020) .

[126] Gleixner, S., Demissie, T. & Diro, G. T. Did ERA5 improve temperature and precipitation reanalysis over East Africa? *Atmosphere* **11** (9), 996 (2020) .

[127] Graham, R. M., Hudson, S. R. & Maturilli, M. Improved performance of ERA5 in arctic gateway relative to four global atmospheric reanalyses. *Geophysical Research Letters* **46** (11), 6138–6147 (2019) .

[128] Gelaro, R. *et al.* The modern-era retrospective analysis for research and applications, version 2 (merra-2). *Journal of climate* **30** (14), 5419–5454 (2017) .

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

[129] Oceanic, N. & (NOAA), A. A. Hourly/sub-hourly observational data. website. URL https://www.ncei.noaa.gov/maps/hourly/n. Accessed: 2023-12-28.

[130] Van Ruijven, B. J., De Cian, E. & Sue Wing, I. Amplification of future energy demand growth due to climate change. *Nature communications* **10** (1), 2762 (2019) .

[131] Dirks, J. A. *et al.* Impacts of climate change on energy consumption and peak demand in buildings: A detailed regional approach. *Energy* **79**, 20–32 (2015) .

[132] Craig, M. T. *et al.* Overcoming the disconnect between energy system and climate modeling. *Joule* **6** (7), 1405–1417 (2022) .

[133] Jentsch, M. F., James, P. A., Bourikas, L. & Bahaj, A. S. Transforming existing weather data for worldwide locations to enable energy and building performance simulation under future climates. *Renewable energy* **55**, 514–524 (2013) .

[134] Machard, A., Inard, C., Alessandrini, J.-M., Pelé, C. & Ribéron, J. A methodology for assembling future weather files including heatwaves for building thermal simulations from the European coordinated regional downscaling experiment (EURO-CORDEX) climate data. *Energies* **13** (13), 3424 (2020) .

[135] O'Neill, B. C. *et al.* The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development* **9** (9), 3461–3482 (2016) .

[136] May, R. M. *et al.* MetPy: A Meteorological Python Library for Data Analysis and Visualization. *Bulletin of the American Meteorological Society* **103** (10), E2273 – E2284 (2022). URL https://journals.ametsoc.org/view/journals/bams/103/10/BAMS-D-21-0125.1.xml. https://doi.org/10.1175/BAMS-D-21-0125.1 .

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

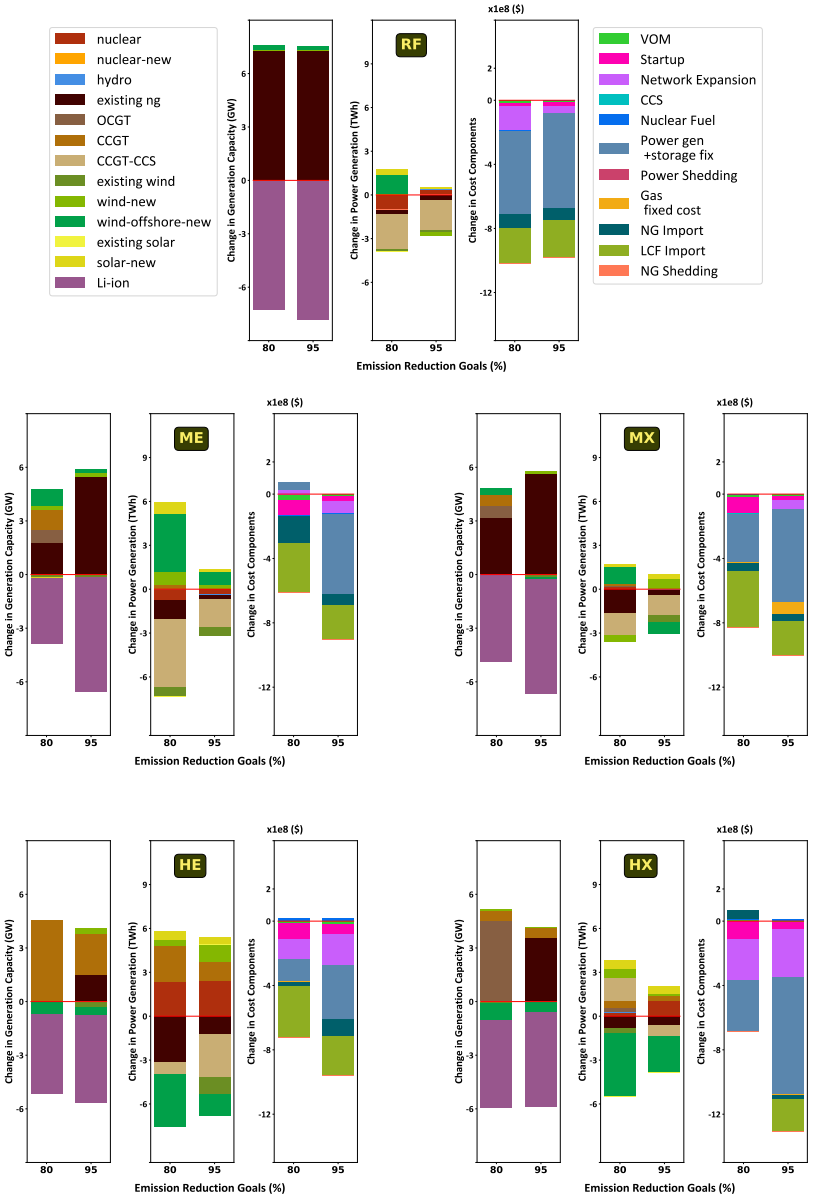
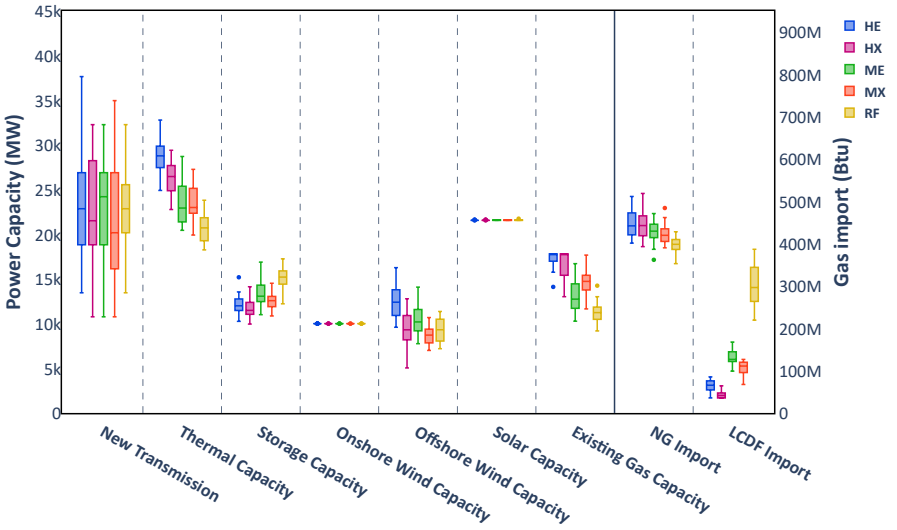
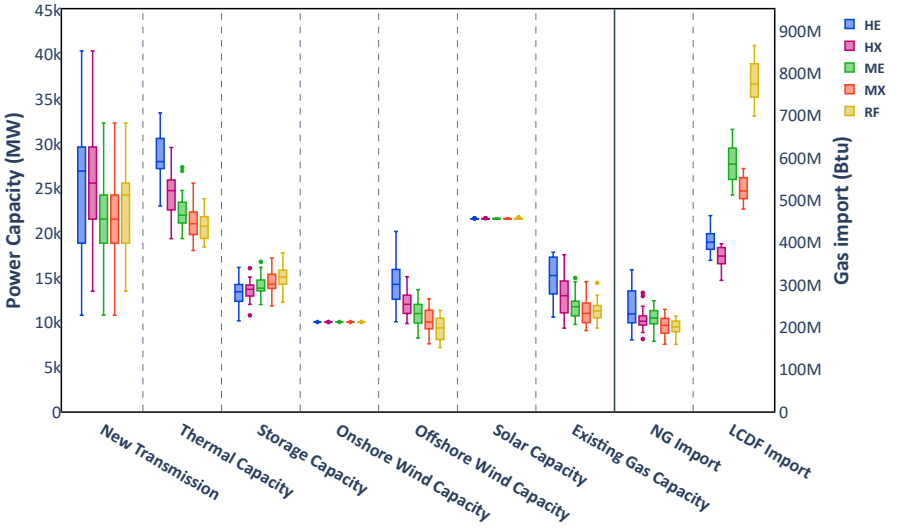


Fig. S10: Difference in capacity, generation, and cost components between instances with and without transportation flexibility for exemplar weather year 2003 under different decarbonization targets and electrification scenarios.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.



(a) 80% decarbonization target



(b) 95% decarbonization target

Fig. S11: Ranges of capacity and gas consumption metrics for all 20 weather years under each emissions constraint. "Thermal capacity" includes both new and existing gas assets. Unless otherwise specified, all capacity metrics include both existing brownfield capacity and new assets.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

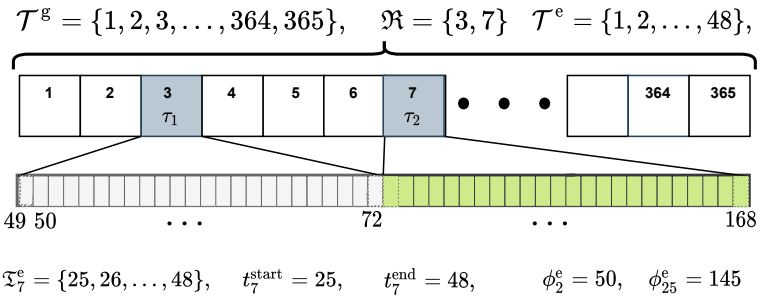


Fig. S12: Illustration of the model's temporal resolution of two representative days used for power systems operations. The top row represents planning days in a year, with days 3 and 7 as representative days forming the set $\mathfrak{R}$. The bottom row shows the hours corresponding to representative days. The set of planning periods for the gas system ($\mathcal{T}^{\mathrm{g}}$) is the entire year. The set of planning periods for the power system ($\mathcal{T}^{\mathrm{e}}$) is the chronologically ordered set of hours in the representative days whose mapping to their original index is given by $\phi^{\mathrm{e}}_t, t \in \mathcal{T}^{\mathrm{e}}$. The set of hours in their original indexing is denoted by $\mathfrak{T}^{\mathrm{e}}_\tau$ with t^{start}_τ and t^{end}_τ signifying the starting and ending hours for the representative day τ. The representative days are chronologically ordered but are not necessarily contiguous.

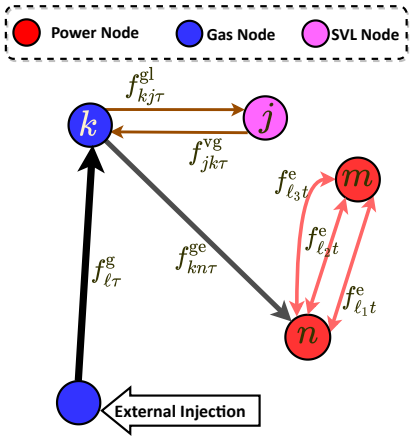


Fig. S13: Flow variables between different nodes. In our model, power nodes can be connected by multiple bi-directional transmission lines denoted by $f^{\mathrm{e}}_{\ell t}$. Each power node operates a set of local gas-fired plants by drawing gas from its closest gas node. The variable $f^{\mathrm{ge}}_{kn\tau}$ captures this flow. Each gas node is connected to its closest SVL node through two unidirectional pipelines where one is from gas to SVL's liquefaction facilities denoted by $f^{\mathrm{gl}}_{kj\tau}$; and the other one from SVL's vaporization facility to gas node denoted by $f^{\mathrm{vg}}_{jk\tau}$. The variable $f^{\mathrm{g}}_{\ell\tau}$ denotes the pipe flow between gas nodes. gas 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.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

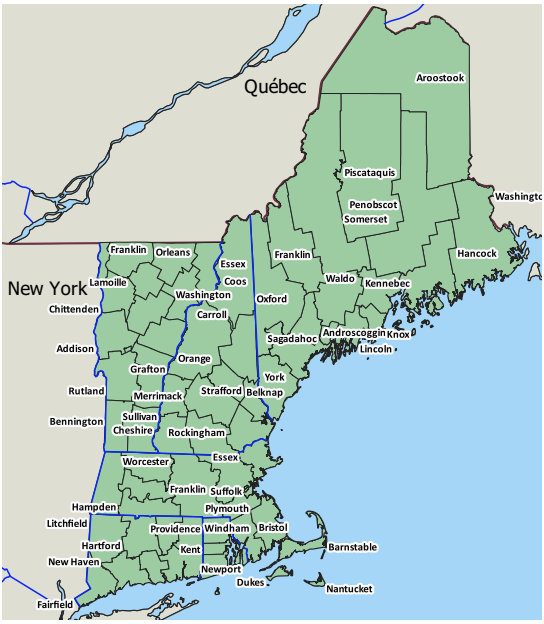


Fig. S14: New England counties

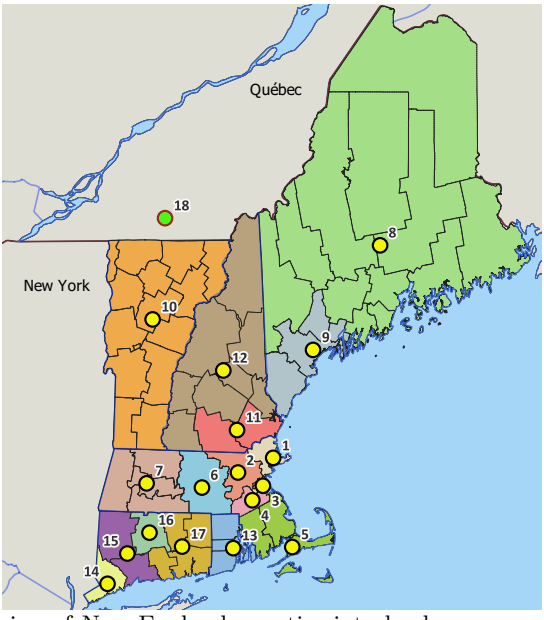


Fig. S15: Grouping of New England counties into load zones, and the location of power nodes including the import node in Quebec

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

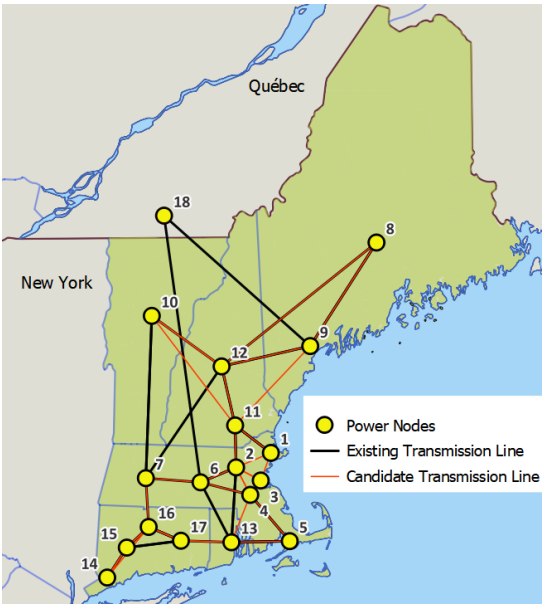


Fig. S16: Existing and candidate transmission lines in the power network

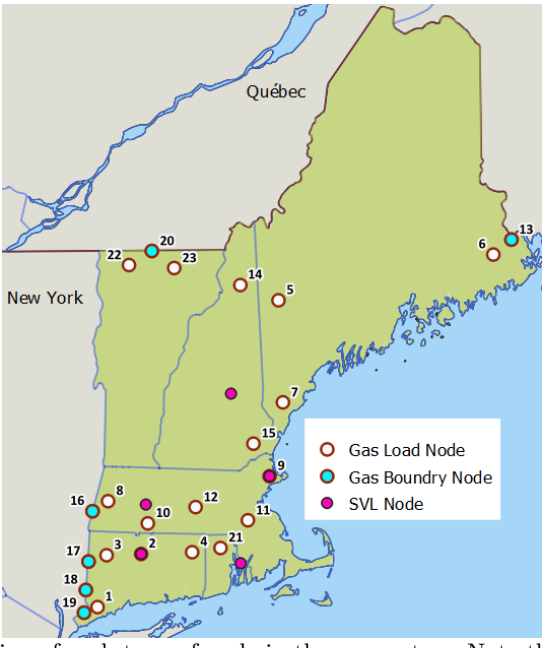


Fig. S17: Location of each type of node in the gas system. Note that two SVL and load nodes share the same location. SVL nodes are not numbered here for clarity of the figure.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

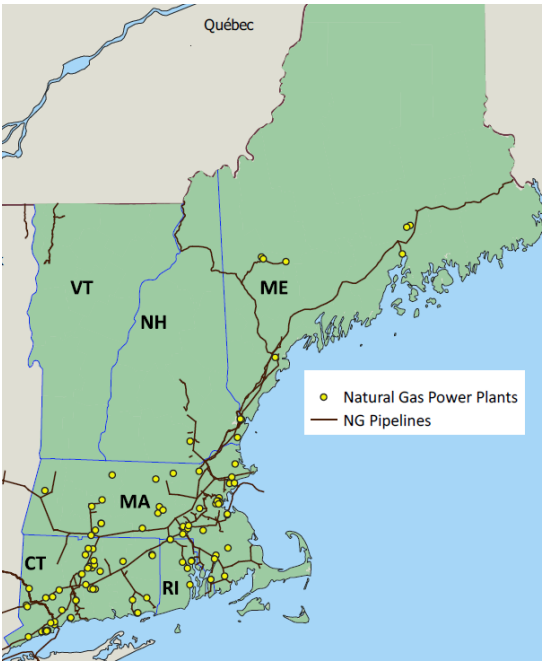


Fig. S18: Existing NG pipelines and gas-fired power plants in New England

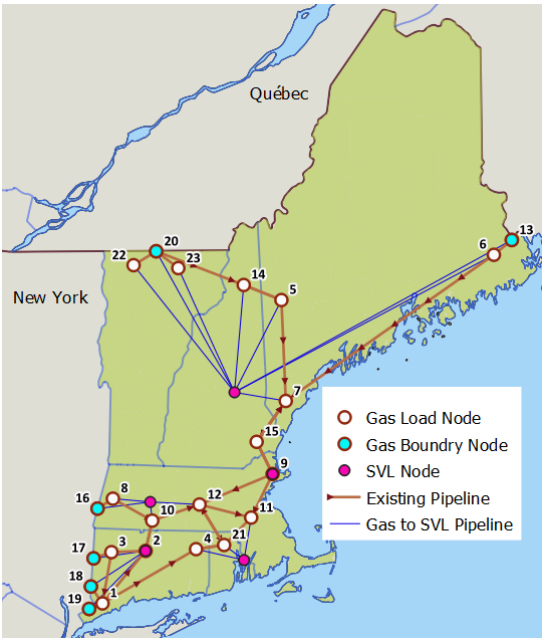


Fig. S19: Existing pipelines between three types of gas nodes. The connections between load and SVL nodes are shown by a single line for clarity of the figure.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

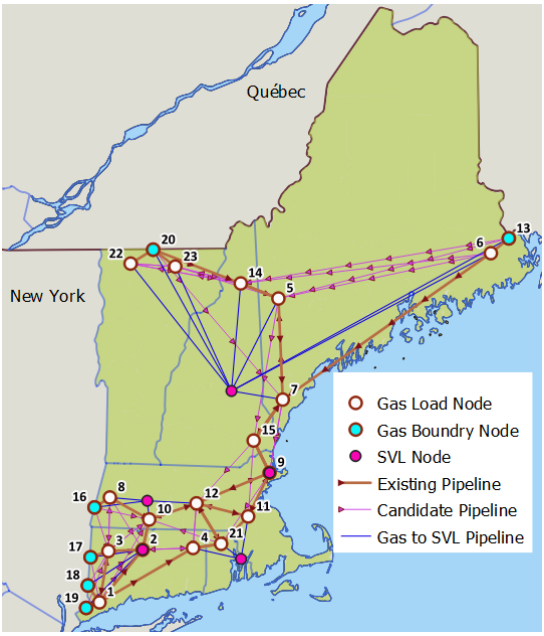


Fig. S20: Existing and candidate pipelines between three types of gas nodes

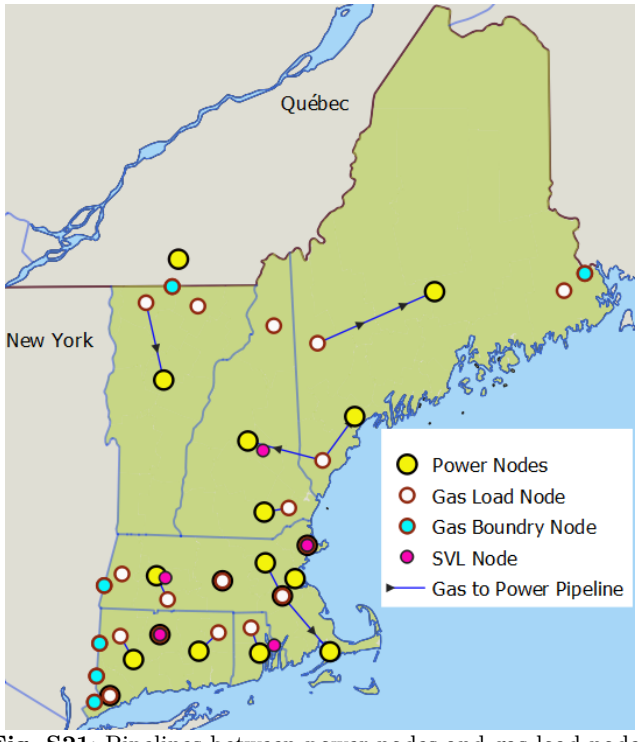


Fig. S21: Pipelines between power nodes and gas load nodes

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

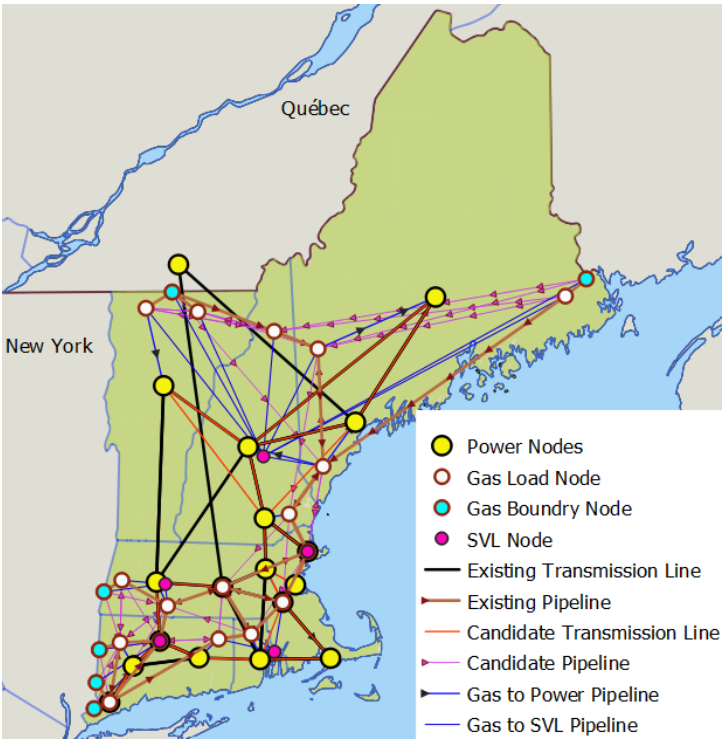


Fig. S22: All existing and candidate connections in New England case study

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

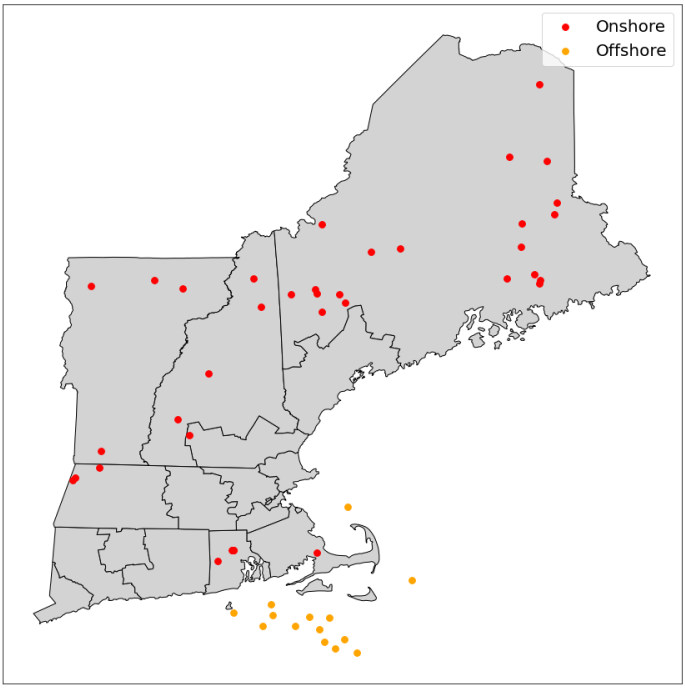


Fig. S23: Map of onshore and offshore wind locations used to calculate zone-level capacity factors for weather years 2001-2020.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

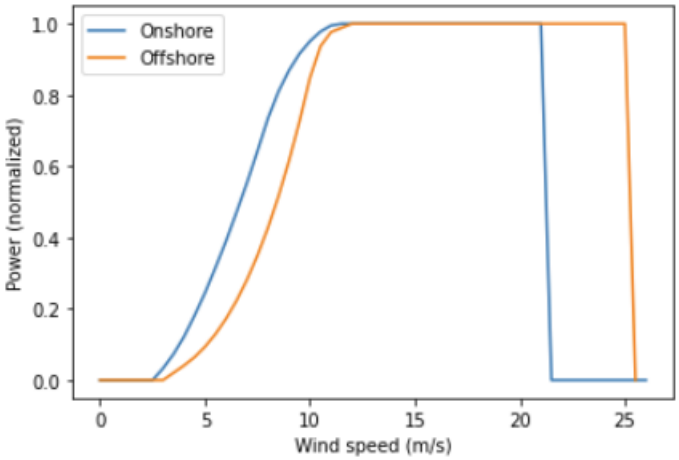


Fig. S24: Power curve for onshore wind comes from data for the Gamesa:G126/2500 turbine as presented in Brown and Botterud [34]. Power curve for offshore wind is taken from the NREL Reference 6MW offshore turbine model proposed in [109]

.

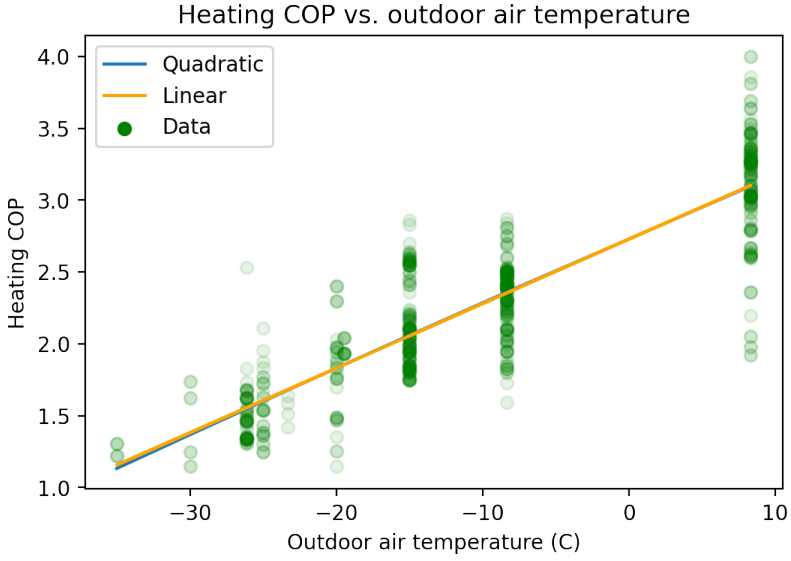


Fig. S25: Regression models of the NEEP data for heating COP versus outdoor air temperature. We chose the linear model for our analysis. The model extrapolates beyond the range of the data displayed.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

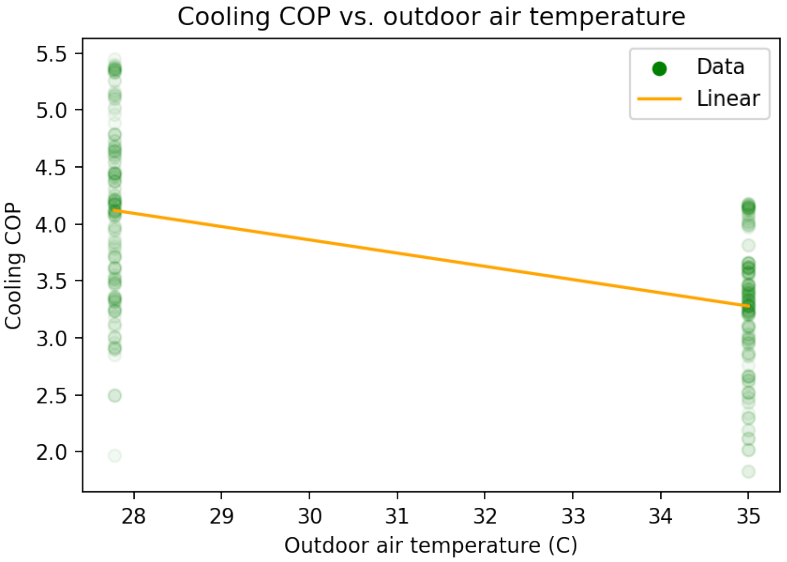


Fig. S26: Regression models of the NEEP data for cooling COP versus outdoor air temperature. The model extrapolates beyond the range of the data displayed.

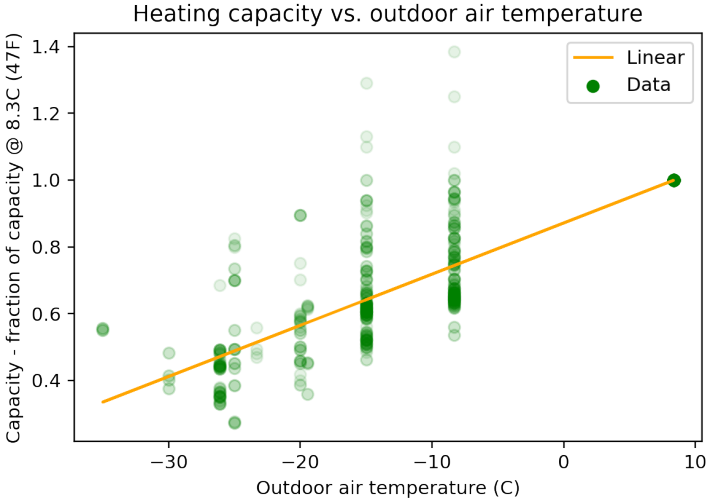


Fig. S27: Regression model of the NEEP data for heating capacity versus outdoor air temperature. Because the heat pumps in our model can have different capacities, we normalize the curve to the capacity at 8.3^{o}C (47^{o}F). The model extrapolates beyond the range of the data displayed.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

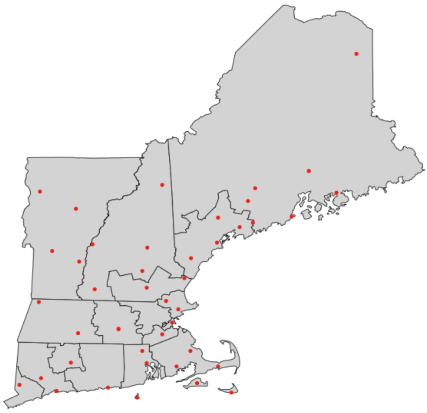


Fig. S28: Map of locations of weather data used in the analysis.

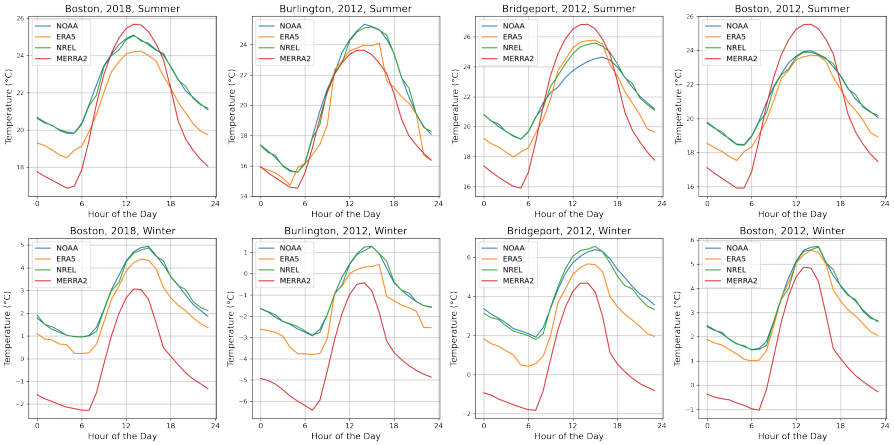


Fig. S29: Comparison of different weather dataset temperature series against NOAA observational data for varying locations and periods. Profiles show average temperatures at each hour of the day for the season indicated. NOAA data from [129], NREL data from [25], MERRA2 data from [128]

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.

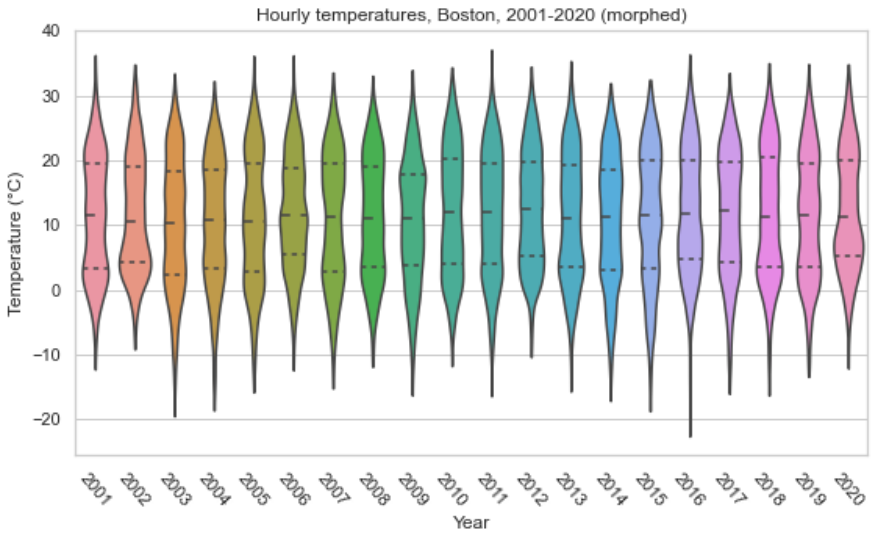


Fig. S30: Temperatures for Boston across the 20 weather years, including the morphing to approximate climate change in 2050. The width of the violin is proportional to the number of hours at the given temperature. The half-line represents the median, and the dashed lines represent the quartiles. The extremes indicate the maximum and minimum annual temperatures.

Pre-publication paper for Future Energy Systems Center sponsor companies only. Not for public distribution.