**Nominations**To nominate a project, please complete [this form](https://docs.google.com/forms/d/e/1FAIpQLSehL79tSri7dvW9lBty2C3krK4utlnfyHfc6YIVzXoY9V7EsQ/viewform?usp=sf_link) by September 15 at 5pm. As part of the nomination process, you will be asked to upload a 1-2 paragraph (up to 600 words) description of the project. This should include:

* A description of the context and impact of the project and the relevance of open data, in language broadly accessible to people from other disciplines. Links or figures can be included.
* The role that the nominee(s) played in the project.

**JPoNG Data**: high-fidelity supply/demand projections and network topology data for decarbonized power and natural gas planning

The decarbonizing transition from the current fossil-dominated energy system to renewable resources requires coordinated planning of future electric power and natural gas (NG) systems, which are some of the largest and most sophisticated infrastructural systems globally. In this regard, forthcoming policy decisions for designing society’s decarbonized power-gas systems crucially depends on high-fidelity modeling of the energy system as well as credible energy supply/demand projections. This is particularly the case in a future context where the electrification of residential heating systems will result in electricity demand that is far more sensitive to extreme weather events. The impact of heating electrification on energy demand patterns is underexplored relative to other sectors such as electric vehicles. However, its projection has the potential to significantly jeopardize the stability of energy systems globally. For example, it can overwhelm the current power supply capacity or render the gas system defunct from disuse. Yet, there is no robust projection methods to produce high-validity planning outcome to inform utility companies and policymakers.

Owing to this importance, we offer the first-known dataset that enables demand-supply mutual planning in the context of a particularly challenging cold-climate case study. **Da**taset for the **s**ustainable **e**nergy **t**ransition under **h**eating electrification (DASETH) provides a comprehensive data repository that enables the modeling of power-gas system futures under deep heating electrification. The dataset considers 6,17 and 67 subregions in New England and contains 1) hourly projection of power demand and natural demand for the year 2050 under 20 different weather years and 5 different heating electrification levels (100 projections), as shown in Fig. 1; 2) hourly projection of availability factors for solar, wind, and offshore wind generators for the year 2050 under 20 different weather years and 5 different heating electrification levels (100 projections); 3) detailed topological description of the current and future power and natural gas infrastructure in the region load zones, transmission lines and pipelines, as depicted in Fig. 2; 4) other parameters for energy planning including parameters for possible power plants and storage technologies. The detail description for constructing each dataset is provided in [1].

Despite its recent development, the dataset is already used in several publications [1,2,3,4,5] and is being used in continuing work by three different groups between MIT and Princeton University, and we plan to continue promoting its dissemination. While the dataset is currently curated for New England region, the application of the dataset to this case study has already enabled us to draw some novel policy-relevant conclusions for future energy system planning. For example, we find that the sensitivity of heating demand to extreme cold weather events in a decarbonized future calls for a diverse energy resource mix with deep deployment of renewables and novel technologies, and that pursuing aggressive electrification (and reductions in NG consumption) along with energy efficiency measures will result in the least-cost energy system outcomes [1]. Furthermore, our group is working to extend the dataset to a much wider range of states across the Midwest and Mid-Atlantic regions, enabling policy-relevant analyses by stakeholders across broad swathes of the country.

Role of nominees: Rahman Khorramfar was the lead person to develop energy topology and related parameters. Morgan Santoni-Colvin has developed power/gas load and availability factors for renewables.

A diagram of a diagram

Description automatically generated

**Fig. 1: Three step procedure to project load data across 20 weather years and 5 electrification scenarios**

A map of the state of new york

Description automatically generated

**Fig. 2: Topology of power and NG infrastructure in New England. The topology includes the existing infrastructure and the possible future developments.**

References:

1. Khorramfar, Rahman, Morgan Santoni-Colvin, Saurabh Amin, Leslie K. Norford, Audun Botterud, and Dharik Mallapragada. "Cost-effective Planning of Decarbonized Power-Gas Infrastructure to Meet the Challenges of Heating Electrification." arXiv preprint arXiv:2308.16814 (2023).
2. Khorramfar, Rahman, Dharik Mallapragada, and Saurabh Amin. "Electric-Gas Infrastructure Planning for Deep Decarbonization of Energy Systems." *arXiv preprint arXiv:2212.13655* (2022).
3. Khorramfar, R., Santoni-Colvin, M., Mallapragada, D., Amin, S., Norford, L. K., & Botterud, A. (2022, December). Decarbonized Future Power-Gas Systems under Large-Scale Electrification of Building Heating. In *AGU Fall Meeting Abstracts* (Vol. 2022, pp. GC42Q-0922).
4. Brenner, Aron, Rahman Khorramfar, and Saurabh Amin. "Learning Spatio-Temporal Aggregations for Large-Scale Capacity Expansion Problems." *Proceedings of the ACM/IEEE 14th International Conference on Cyber-Physical Systems (with CPS-IoT Week 2023)*. 2023.
5. Brenner, A., Khorramfar, R., Mallapragada, D., & Amin, S. (2022). Graph Representation Learning for Energy Demand Data: Application to Joint Energy System Planning under Emissions Constraints. *arXiv preprint arXiv:2209.12035*.

The dataset is constructed for New England regions with varying spatial granularity for the power network. The residential load, however, is available for the power system with 17 nodes and is accessible at folder ‘Raw\_Residential\_Data’.

Insider each of the remaining folder (i.e., ‘6\_Power\_Nodes’, ‘17\_Power\_Nodes’, and ‘67\_Power\_Nodes’) there are two folders:

* Gas\_System\_Data that contain all the data relevant of the natural gas (NG) network in the region. The gas network is remains a 23-node network for all granularity of the power network.
  + ‘NG\_AdjE\_Nodes.csv’: adjacent power nodes for each gas node
  + ‘NG\_Load\_{Elec\_Scenario}\_BaseYear{WY}.csv’: these files contain the daily demand of gas across. The {Elec\_Scenario} keywork takes one of the 5 electrification scenarios (HE, HX, ME, MX, RF) and {WY} takes one of the 20 weather years between 2001 and 2020.
  + NG\_Nodes.csv: relevant information for each gas node including their locations and associated counties.
  + ‘NG2NG\_Pipelines.csv’: existing and candidate pipelines in the gas network
  + ‘SVL\_data.csv’: information for each gas storage-vaporization-liquefaction (SVL) node including their locations and associated counties.
  + ‘SVL\_params.csv’: parameters for SVL nodes
* ‘Power\_System\_Data’
  + ‘AvailabilityFactor\_{VRE}\_BaseYear{WY}.csv’: these files contain the hourly availability factors for renewable generators. The keywork {VRE} takes ‘Solar’, ‘Wind\_Offshore’, and ‘Wind\_Onshore’ values and {WY} takes one of the 20 weather years between 2001 and 2020.
  + ‘Electricity\_Load\_{Elec\_Scenario}\_BaseYear{WY}.csv’: these files contain the hourly demand of electric power. The {Elec\_Scenario} keywork takes one of the 5 electrification scenarios (HE, HX, ME, MX, RF) and {WY} takes one of the 20 weather years between 2001 and 2020.
  + ‘Plant\_params.csv’: parameters for the existing and possible future power generators.
  + ‘Plants\_Nodes’: The number of existing power generators in each power node.
  + ‘Power\_Nodes.csv’: information for each electric power node including their locations and the associated counties.
  + ‘Regional\_multipliers.csv’: the relative cost of establishing renewable generators in each state.
  + ‘Storage\_params.csv’: parameters for battery storage technologies.
  + ‘Transmission\_Lines.csv.’: existing and candidate transmission lines.