A hetero-functional graph structural analysis of the American Multi-Modal Energy System

Dakota J. Thompson [a],[*], Amro M. Farid [b],[c]

[a] *Thayer School of Engineering at Dartmouth, Hanover NH, United States*
[b] *Stevens Institute of Technology, School of Systems & Enterprises, Hoboken, NJ, United States*
[c] *MIT Mechanical Engineering, Cambridge MA, United States*

ARTICLE INFO

ABSTRACT

As one of the most pressing challenges of the 21st century, global climate change demands a host of changes across four critical energy infrastructures: the electric grid, the natural gas system, the oil system, and the coal system. Unfortunately, these four systems are often studied individually, and rarely together as integrated systems. Instead, holistic multi-energy system models can serve to improve the understanding of these interdependent systems and guide policies that shape the systems as they evolve into the future. The NSF project entitled "American Multi-Modal Energy System Synthetic & Simulated Data (AMES-3D)" seeks to fill this void with an open-source, physically-informed, structural, and behavioral machine-learning model of the AMES. To that end, this paper uses a GIS-data-driven, model-based system engineering approach to develop structural models of the American Multi-Modal Energy System (AMES). This paper produces and reports the hetero-functional incidence tensor, hetero-functional adjacency matrix, and the formal graph adjacency matrix in terms of their statistics. This work compares these four hetero-functional graph models across the states of New York (NY), California (CA), Texas (TX), and the United States of America (USA) as a whole. From the reported statistics, the paper finds that the geography and the sustainable energy policies of these states are deeply reflected in the structure of their multi-energy infrastructure systems and impact the full USA structure.

1. Introduction

As one of the most pressing challenges of the 21st century, global climate change demands a host of changes across at least four critical energy infrastructures: the electric grid, the natural gas system, the oil system, and the coal system. In the context of the United States, this paper refers to this system-of-systems as "the American Multi-Modal Energy System (AMES)". The needs of climate change demand mitigation and adaptation strategies which are far more demanding than the needs of just mitigation alone. Therefore, as policies are developed to drive the sustainable energy transition forward, they must not just aim to *mitigate* climate change but must also *adapt* to its effects with resilient architectures. In effect, the need for decarbonization must be harmonized with the need for economic development, national energy security, and equitable energy access [1–3]. These combined requirements to develop effective policies necessitate an understanding of the AMES interdependencies and how they vary geographically and temporally [4]. Furthermore, this cross-sectoral interdependency can introduce architectural fragility [5] that must be managed as an integral part of the sustainable energy transition.

Holistic multi-energy system models should serve to improve the understanding of these interdependent systems as they evolve into the future. Unfortunately, while there have been many attempts at modeling multi-energy systems and large-scale flows of energy, the field remains relatively nascent [6–14]. While initial models were developed in response to the oil crisis and then to mitigate climate change with decarbonization they have also been developed to introduce new energy streams such as hydrogen, synthetic fuels, bio-fuels, and other renewable energy sources. These works introduce their own set of weaknesses including the lack of asset level granularity, difficulty of use, and specific one-off geographically-specific use case models. Additionally, a majority of the works investigating these energy systems in the past have been performed on individual energy networks [15–19]. More recently, work has been published analyzing only a couple of systems together such as pairing the electric grid with one of the other fossil fuel systems that compose the AMES [6–13]. These works, however, do not include all four critical energy infrastructures, their many energy commodities, and do not extend to the entire American geography. The complexity of integrating multiple infrastructures

---

scales as $n^2 * m$ where n is the number of infrastructures and m is the number of energy commodities and increasingly requires robust model-based systems engineering and graph theoretic methods to manage these many interfaces. As an exception, the EIA developed a comprehensive model called the National Energy Modeling System (NEMS) which it uses to produce the (American) Annual Energy Outlook [20]. Despite serving this important function and being publicly available, this software tool remains opaque and difficult to use. The EIA website itself recognizes: "[The] NEMS is only used by a few organizations outside of the EIA. Most people who have requested NEMS in the past have found out that it was too difficult or rigid to use [21]". In addition to the works on pairing the electric grid with individual fossil fuels, extensive literature has developed on the co-dependence of the electric grid and water resources in the form of the Energy-Water Nexus (EWN) [22–41]. These works have demonstrated the potential benefits of analyzing and optimizing the co-dependent infrastructure systems as a single system-of-systems rather than each system independently [42]. Therefore, seeing the benefits pairing two systems can provide, further emphasizes the need to model the entirety of the AMES. Consequently, holistic multi-energy system models of the AMES remain a present need for open-source, citizen-science to inform policies.

With a deficit of spatially and functionally resolved data, and with the current methods for modeling multi-energy systems having their limitations, the (American) National Science Foundation (NSF) put forth a call for "research to develop and make available simulated and synthetic data on interdependent critical infrastructures (ICIs), and thus to improve understanding and performance of these systems" [43]. The NSF project entitled "American Multi-Modal Energy System Synthetic & Simulated Data (AMES-3D)" seeks to fill this void with an open-source structural and behavioral model of the AMES. Adhering to a Model-Based Systems Engineering (MBSE) approach [44,45], this project develops an interdependent system data set and its associated models on top of a strong theoretical foundation in systems engineering. As a result, it can be used for practical applications in the energy systems field to address not just mitigation of climate change but adaptation and resilience as well. This is made possible by using asset-level, openly available datasets to infer the AMES' reference architecture [46] to organize and define the interconnections between the four subsystems. The reference architecture uses SysML [45] to model the four interdependent energy systems, and the flows of mass and energy within and between them. This reference architecture provides a more detailed and self-consistent MBSE foundation for energy models moving forward relative to the "reference energy systems" that have been used in some national energy system optimization models [47–50]. The datasets that infer the reference architecture are also used to instantiate the AMES into an instantiated architecture [46]. While the NSF project seeks to develop both a structural and behavioral model, this paper focuses its scope on the former.

The development of the AMES reference architecture provides several immediate benefits. The first is that a SysML-based reference architecture describes the system's form, function, and the allocation of the latter onto the former [44,45]. Therefore, the reference architecture describes not just what the system is made of, but also what it does. Second, a SysML-based reference architecture can be readily translated into mathematical models including both the form and function. Standard structural models include formal graphs that describe energy facilities and how they are connected. In the meantime, hetero-functional graphs (HFG) describe how the wide variety of capabilities in a system are interconnected and the flows of operands between them. HFGs have been shown to provide more information than formal graphs when analyzing an evolving instantiated architecture [51,52]. In effect, HFGT provides a means to quantitatively interpret the graphical SysML-based models from both a formal and functional lens. Such an analysis has already been conducted on a small-scale electric power distribution systems [51] as well as on a large scale for the entirety of the American electric power system [52]. This paper now builds on these electricity-only analyses to study the structure of the whole AMES.

1.1. Original contributions

This paper uses a data-driven, MBSE-guided approach to develop open-source structural models of the American Multi-Modal Energy System. More specifically, the AMES reference architecture [46] is applied to an asset-level GIS data called Platts Map Data Pro [53] to create models of several regions. The instantiated structural models include for the first time the electric grid, the natural gas system, the oil system, the coal system, and the interconnections between them as defined by the AMES reference architecture, for the full contiguous United States of America (USA). Initial results are organized into two categories; a formal and hetero-functional graph for each of the regions being studied, New York (NY), California (CA), Texas (TX), and the full contiguous (USA). The states are chosen for their large size and the diversity of their energy policies. Consequently, the chosen regions have also taken distinct directions to advance the sustainable energy transition. In 2019, CA had the most renewable energy generation out of all the states [54]. In the meantime, NY's efforts to expand renewable energy capacity are balanced by its reliance on natural gas and oil to meet space heating energy demands [55]. Alternatively, Texas, while being the nation's leading crude oil and natural gas producing state, is also the nation's leading producer of wind-powered electric generation [56]. By using MBSE and HFGT, new open-source data models are presented for these three states and the full USA to aid in advancing and guiding the sustainable energy transition and energy policies.

1.2. Paper outline

The remainder of the paper proceeds as follows. Section 2 is a description of the background literature and the lack of open data models used to develop the instantiated architecture models and guide policy. Section 3 then presents the data source that drove the instantiated models in Section 3.1, followed by defining the AMES Reference Architecture in Section 3.2. The data processing in Section 3.2 is then presented followed by a subsection on hetero-functional graph theory 3.4. The paper then presents a comparison of the formal graphs and hetero-functional graphs network statistics for each state in the Results Section 4. The Results first start with an analysis of the computational intensity in Section 4.1, it then presents the formal and hetero-functional graph statistics in Sections 4.2 and 4.3 respectfully. Section 4 continues to present the formal graph resource distributions in Section 4.4 and the HFG capability distributions in Section 4.5. A discussion of the HFG process degree distributions in Section 4.6 brings the methodology section to a close. Finally the paper is brought to a conclusion in Section 5.

2. Background

This section serves to situate the paper with respect to existing literature and the ongoing trends in the field. Section 2.1 describes existing multi-energy system models. Section 2.2 discusses the general lack of interdependent infrastructure system data and models. Section 2.3 discusses the emerging trend towards open data and source code in the multi-energy system field. Finally, Section 2.4 introduces some essential concepts in HFGT that are used in the remainder of the paper.

2.1. Existing multi-energy system models

Many existing multi-energy system models are in effect national models used to inform national energy policy. Many such national models have an often implicitly stated "reference energy system" which serves as the first step for defining energy flow relationships [47–50].

Definition 1 (*Reference Energy System [47]*)**.** As a tool to begin modeling energy systems, *"it represents a simplified and aggregated graphical representation of the entire energy system under analysis which shows all existing and potential new energy supply chains from primary energy to final demand".* ■

A reference energy system defines how the primary energy flows are processed and converted into different energy carriers and where the energy is ultimately utilized in end-use sectors. While the RES does identify all the modes of energy being tracked through the model, it is driven purely by system behavior. This means it only incorporates the system function but does not explicitly describe the system form. In contrast, the practice of MBSE requires the definition of system function and system form [44,45] in a graphical modeling language like SySML. When a "reference energy system" omits the system form, asset-level, and technology-specific interactions are ultimately neglected. Additionally, by ignoring the form at the starting point for these energy system models, simulation model outputs and results become less transparent. Consequently, the absence of a description of the system form further impedes the comparability of different models because they do not explicitly state which technologies exist and how they interact. This lack of model comparability further impedes the comparability of differing strategies to global climate change mitigation and adaptation. As an alternative to RES models, MBSE and SysML utilize reference architectures as the basis of all downstream modeling activities.

Definition 2 (*Reference Architecture [57]*). *"The reference architecture captures the essence of existing architectures, and the vision of future needs and evolution to provide guidance to assist in developing new instantiated system architectures. …Such reference architecture facilitates a shared understanding across multiple products, organizations, or disciplines about the current architecture and the vision on the future direction. A reference architecture is based on concepts proven in practice. Most often preceding architectures are mined for these proven concepts. For architecture renovation and innovation validation and proof can be based on reference implementations and prototyping. In conclusion, the reference architecture generalizes instantiated system architectures to define an architecture that is generally applicable in a discipline. The reference architecture does however not generalize beyond its discipline".* ∎

A reference architecture specifically includes a description of system form and function. Consequently, it makes transparent all flows of matter and energy and which energy assets are used for transforming and transporting these flows. Furthermore, It becomes much easier to compare energy system models and their underlying assumptions. Finally, it is straightforward to determine how much of a reference architecture appears in the instantiated architecture that pertains to the geography of a specific region or case study.

Definition 3 (*Instantiated Architecture [45,57]*). A case-specific architecture, which represents a real-world scenario, or an example test case. At this level, the physical architecture consists of a set of instantiated resources, and the functional architecture consists of a set of instantiated system processes. The mapping defines which resources perform what processes. ∎

The transparency and comparability of a reference architecture becomes even more valuable in light of the open energy modeling initiative. It began in 2014 as a result of the open access movement which had begun in 2010 [47,58]. These movements originated from a recognition of the limited transparency and reproducibility of energy systems models [59]. The open energy modeling initiative seeks to encourage the use of open access models in research to guide energy policies and the sustainable energy transition. While this initiative has resulted in the production of many open-access models, much work remains to compare and converge these models towards the ultimate goals of sustainable energy transition.

Several powerful energy system models have come out of the energy modeling initiative including the OSeMOSYS, NEMS, and PRIMES [47, 60–62]. However, these models each had their weaknesses. For example, as previously mentioned, the NEMS is extremely difficult to use making it unusable for many open access projects. While the NEMS was a large model incorporating many different modes of energy to inform the annual energy outlook for policy planning, there have also been a plethora of models developed for specific systems rather than focusing on the interconnection of multiple critical infrastructures [15–19]. Alternatively, other energy system optimization models (ESOMs) such as the EnergyScope TD have been designed for the analysis of intermittent integration of renewable on an hourly scale. Similarly, the Electric Power Enterprise Control Simulator addresses the integration of variable renewable energy on multiple time scales [63,64]. These operational time-scale models, however, are not designed to explore the annual transformation pathways of the sustainable energy transition [47,65,66]. Finally, the development of multi-energy system models did not truly gain concerted attention until 2016 [67].

2.2. Interdependent infrastructure system data & models

As the multi-energy system model developed, it created a greater need for interdependent infrastructure system data and models. In 2016, the NSF released a call for open interdependent critical infrastructure system data [43] that could be specifically used in system resilience and climate change adaptation research. This call directly addresses the lack of existing open data sources on energy systems.

In the meantime, the existing methods for organizing such data into multi-modal energy system models had their own limitations. Multi-layer networks, for example, were often looked at as the leading candidate to structurally model these interdependent infrastructures. Unfortunately, multi-layer networks have often been unable to address the explicit heterogeneity often encountered in engineering systems [68,69]. In a recent comprehensive review, Kivela et al. [69] wrote:

> "The study of multi-layer networks … has become extremely popular. Most real and engineered systems include multiple subsystems and layers of connectivity and developing a deep understanding of multi-layer systems necessitates generalizing 'traditional' graph theory. Ignoring such information can yield misleading results, so new tools need to be developed. One can have a lot of fun studying 'bigger and better' versions of the diagnostics, models, and dynamical processes that we know and presumably love — and it is very important to do so but the new 'degrees of freedom' in multi-layer systems also yield new phenomena that cannot occur in single-layer systems. Moreover, the increasing availability of empirical data for fundamentally multi-layer systems amidst the current data deluge also makes it possible to develop and validate increasingly general frameworks for the study of networks.
>
> … Numerous similar ideas have been developed in parallel, and the literature on multi-layer networks has rapidly become extremely messy. Despite a wealth of antecedent ideas in subjects like sociology and engineering, many aspects of the theory of multi-layer networks remain immature, and the rapid onslaught of papers on various types of multilayer networks necessitates an attempt to unify the various disparate threads and to discern their similarities and differences in as precise a manner as possible.
>
> … [The multi-layer network community] has produced an equally immense explosion of disparate terminology, and the lack of consensus (or even generally accepted) set of terminology and mathematical framework for studying is extremely problematic".

In addition to the Kivela et al. review [69], Schoonenberg et al. numerically demonstrated the numerical limitations of multi-layer networks relative to hetero-functional graphs [68]. Similarly, Farid et al. theoretically proved the ontological limitations of multi-layer networks relative to hetero-functional graphs [70].

In this context, the NSF funded the "American Multi-Modal Energy System Synthetic & Simulated Data (AMES-3D)" to develop an open-source energy system data set using modeling methods, and more
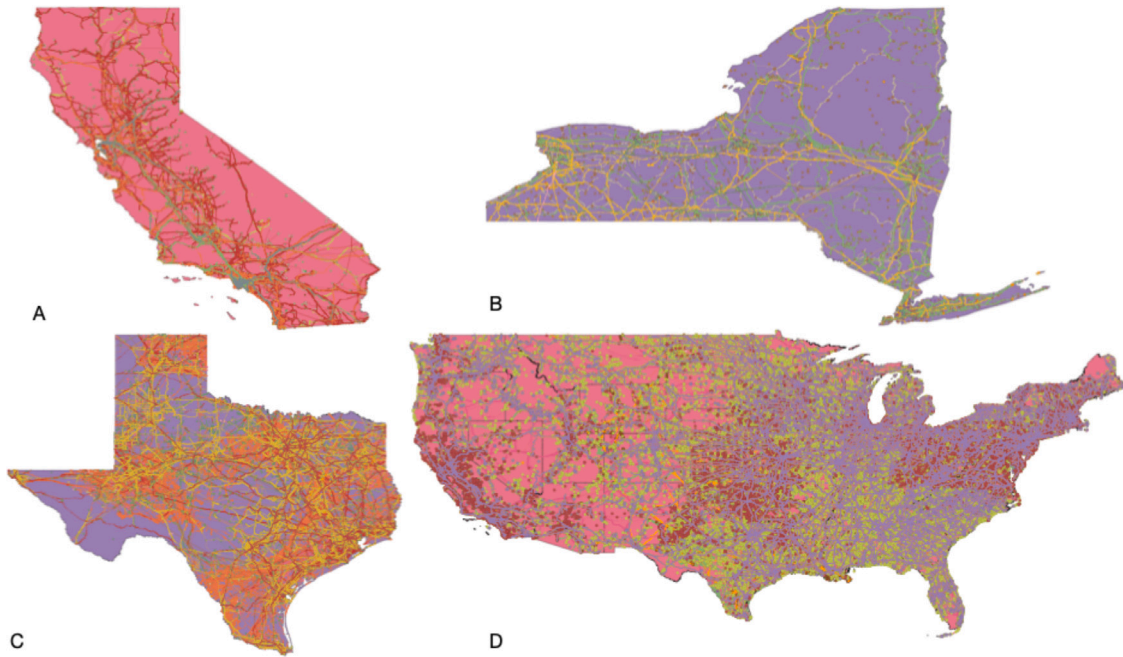
Fig. 1. GIS Layers from the Platts Map Data Pro dataset for the electric grid, natural gas system, oil system, and coal system for California (A), New York (B), Texas (C), and the USA (D).

specifically hetero-functional graph theory (HFGT) that does not exhibit the limitations of multi-layer networks [68]. Consequently, this paper uses HFGT to present a hetero-functional graph structural model (i.e. instantiated architecture) and analysis of the American multi-modal energy system as part of the NSF-funded AMES-3D project.

2.3. Transparency through open data and source code

In addition to the need for reference architectures (Section 2.1), and the need for interdependent infrastructure system data and models (Section 2.2), the multi-energy systems literature also recognizes a pressing need for open data and source code. To that effect, the open modeling initiative (OMI) was founded in 2014 [58]. It was founded with the intent to promote open access models research to guide energy policy [47]. Similarly, the open modeling foundation (OMF) was founded in 2021 [47,71] as an "international open science community that works to enable the next generation modeling of human and natural systems" by making models more easily discoverable and globally accessible [71]. The OMF seeks to create a collection of reusable, interoperable models to study complex interactions between people and the environment at multiple scales.

In direct alignment with the missions of the OMI and OMF is the NSF's call for open interdependent critical infrastructure system data as previously mentioned. The work presented in this paper addresses these needs with transparent open-source code and reproducible mathematics. First, the work relies on the previously developed AMES reference architecture [46] that serves to visualize all of the AMES components, functions, and interactions. In many fields, "mature" reference architectures often develop into international standards that converge and reconcile the otherwise inevitable proliferation of models. Second, the mathematics of hetero-functional graph theory is used to instantiate the AMES reference architecture for NY, CA, TX, and the contiguous United States. The explicit statement of the HFGT mathematics not only validates the work and makes it entirely reproducible, but also makes transparent the open-source HFGT toolbox used to produce the computational results of this work. In all, the open data and open source commitments made in the AMES-3D project have necessitated a research methodology that enhances usability, re-usability, transparency, and comparability of the models, tools, and results.

2.4. Hetero-functional graph theory

In the context of this work, hetero-functional graph theory serves to translate and instantiate a graphical, SysML-based reference architecture into its mathematically equivalent hetero-functional graph. The reliance on a reference architecture grounds the work in a strong MBSE foundation. Additionally, the AMES reference architecture itself provides a clear definition of the assets, functionality, and modes of energy that are included in the AMES. This lossless translation from an MBSE-SysML model to a HFG maintains the allocation of function onto form as capabilities composed of subject + verb + object sentences where the subjects are resources, the predicates are the processes, and the operands are the objects of the verbs. This translation from the AMES reference architecture to a HFG (using the HFGT toolbox) has been well demonstrated for the single-commodity American electric power system [52]. Now this demonstrates the same process to create a multi-energy instantiated model of the full contiguous USA for a structural (network) analysis.

The following HFGT definitions are presented to facilitate the remainder of the paper.

Definition 4 (*System Operand [72]*). An asset or object $l_i \in L$ that is operated on or consumed during the execution of a process. ■

Definition 5 (*System Process [72,73]*). An activity $p \in P$ that transforms a predefined set of input operands into a predefined set of outputs. ■

Definition 6 (*System Resource [72]*). An asset or object $r_v \in R$ that is utilized during the execution of a process. ■

Definition 7 (*Buffer [68]*). A resource $r \in R$ is a buffer $b_s \in B_S$ iff it is capable of storing one or more operands at a unique location in space. $B_S = M \cup B$. ■

Definition 8 (*Capability [74–79]*). An action $e_{wv} \in \mathcal{E}_S$ (in the SysML sense) defined by a system process $p_w \in P$ being executed by a resource $r_v \in R$. It constitutes a subject + verb + operand sentence of the form: "Resource r_v does process p_w". ■
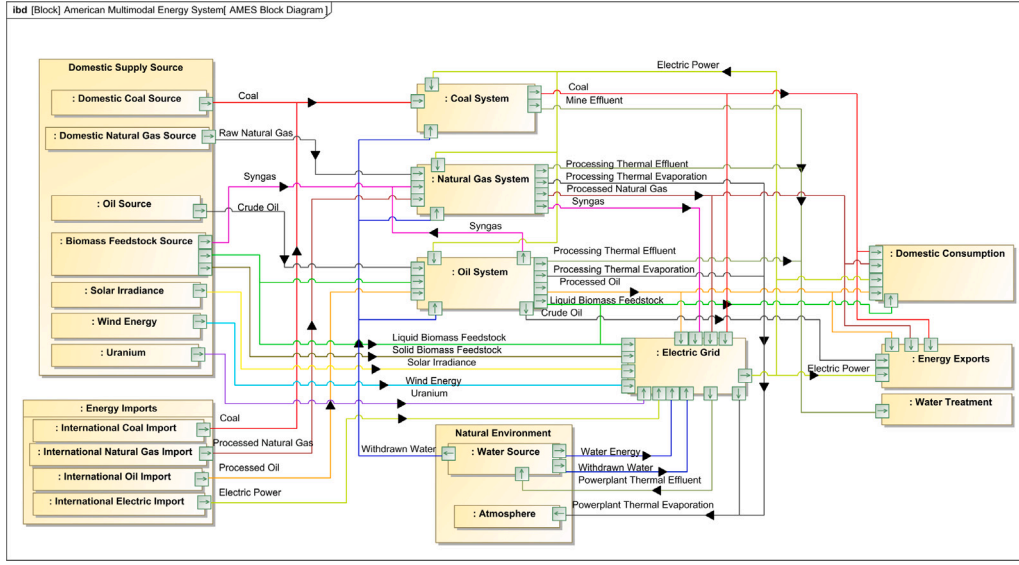
Fig. 2. The top-level internal block diagram of the AMES. The domestic supply sources, energy imports, natural environment, domestic consumption, energy exports, and water treatment are external to the AMES four subsystems of coal, natural gas, oil, and electric grid.

3. Methodology

As mentioned in the introduction, this paper utilizes a data-driven, MBSE-guided methodology to infer a hetero-functional graph structural model of the American Multi-modal Energy System. This section succinctly relays this method in four subsections:

(1) Input Data: Platts Map Data Pro
(2) Infer AMES Reference Architecture
(3) Convert GIS Shape Files to XML Data
(4) Calculate Hetero-functional Graph Structural Model

3.1. Input data: Platts map data pro

Following a data-driven approach, the Platts Map Data Pro [53] data set is used to infer and instantiate the AMES Structural models. This input dataset consists of Graphic Information System (GIS) layers for each of the four subsystems in the AMES [53]. These geo-spatial layers include meta-data attributes of the physical resources/facilities that compose the AMES infrastructure in addition to their GPS co-ordinates. As the Platts Map Data Pro is directed towards wholesale energy decisions, the data is limited to transmission system resources and neglects distribution level assets. This limitation in the dataset notably excludes retail distribution of oil and gas (by truck). It also cuts the electric grid off at the substations excluding distributed electric generation assets such as roof-top solar that are an integral part of the sustainable energy transition. Nevertheless, the Platts Map Data Pro is likely the best available dataset because it allows inferences of not just the AMES's form but its function too. Visualizations of the Platts map data pro GIS layers for each state addressed in this paper and the full USA are presented in Fig. 1.

3.2. Infer AMES reference architecture

The inference of a hetero-functional graph structural model of the American Multi-modal Energy System first requires the inference of the underlying reference architecture. In the data-driven steps that follow, the reference architecture plays a critical role in organizing the cleaned Platts GIS data into defined resources with properly allocated functionality. Fig. 2 shows the top-level context diagram of the AMES reference architecture and it is further elaborated in prior work [46]. The AMES reference architecture provides a consistent blueprint from which to develop AMES models irrespective of the choice of region or scale. It also defines all energy resources/facilities, and the functions that they can perform. The set of operands used to track the flows of mass and energy between its many functions is also defined. Consequently, the AMES reference architecture supports downstream analysis that addresses both climate change mitigation and adaptation.

3.3. Convert GIS shape files to XML data

In this step, the Platts Map Data Pro GIS shape files are converted into an associated XML file for each region (i.e. NY, CA, TX, USA) that serves as the input for the openly available HFGT toolbox [80]. Although the Platts Map Data Pro GIS shape files are a commercially curated dataset, they still need substantial cleaning and processing before being organized into an XML file. As an immediate first step, all resources marked with a canceled status, closed status, or illegible identifying attributes are removed. From there, the primary obstacle is that the Platts GIS data does not guarantee physical continuity between all of the point-type (e.g. refineries and electric power generation facilities) and line-type (e.g. oil pipelines and electric power lines) resources in the dataset.

A novel operand-guided geographical clustering algorithm is developed and then applied to all four energy subsystems simultaneously. It ensures that point-type resources with overlapping GIS coordinates are aggregated into a single resource cluster. It also ensures that line-type resources whose endpoints are not connected to any resource cluster (e.g. "power lines to nowhere") are connected to a nearby resource cluster making sure to maintain the operand compatibility of the line-type resource to the point-type resource cluster. For example, an oil pipeline must connect to a refinery rather than an electric power substation. The clustering algorithm also removes any isolated point resources from the dataset. The geographical clustering algorithm provides an automated means for cleaning the Platts GIS data at the formidable scale presented by the AMES.

The clustering algorithm proceeds in several steps. (1) Using a primary clustering distance of 0.1 miles, point-type resources and the endpoints of line-type resources endpoints with like operand associations are sorted into clusters and connected in a manner that respects the physical continuity of operands. (2) Then with a secondary clustering distance of 1 mile, the remaining isolated nodes are added to existing clusters with like operands. (3) Finally, with a tertiary clustering distance of 35 miles, the remaining isolated nodes are connected to

Table 1
The computational complexity of processing GIS data and the HFGT Toolbox for the Multi-Modal Energy Systems in NY, CA, TX, and the USA.

	New York	California	Texas	United States
Time to assign clusters (s)	2.44	9.9	166.49	3963.55
Time to process data (s)	23.60	105.31	1445.81	49,373.63
Time to process data & Write XML (s)	28.29	121.43	1499.93	51,470.68
Time to calculate A_ρ (s)	14.49	20.25	1252.30	6171.14
Time to run HFGT toolbox (s)	19.13	29.45	1382.97	6439.74
XML size (MB)	5.6	13	192	420.7
HFG resulting pickle size (MB)	8.8	20.6	314.9	682.6

existing clusters with like operands via the creation of new transportation resources. After the execution of this clustering algorithm, the data attributes of these resources are converted to strings in a manner that adheres to the AMES reference architecture and the XML file format required by the hetero-functional graph theory toolbox [80].

3.4. Calculation of hetero-functional graph structural model

The next step is to run the HFGT toolbox [80] on this newly produced XML file so as to produce the positive and negative hetero-functional incidence tensors (HFITs).

Definition 9 (*The Negative 3rd Order Hetero-functional Incidence Tensor (HFIT) $\widetilde{\mathcal{M}}_\rho^-$ [81,82]*). The negative hetero-functional incidence tensor $\widetilde{\mathcal{M}}_\rho^- \in \{0,1\}^{|L| \times |B_S| \times |\mathcal{E}_S|}$ is a third-order tensor whose element $\widetilde{\mathcal{M}}_\rho^-(i, y, \psi) = 1$ when the system capability $\epsilon_\psi \in \mathcal{E}_S$ pulls operand $l_i \in L$ from buffer $b_{s_y} \in B_S$. ∎

Definition 10 (*The Positive 3rd Order Hetero-functional Incidence Tensor (HFIT) $\widetilde{\mathcal{M}}_\rho^+$ [81,82]*). The positive hetero-functional incidence tensor $\widetilde{\mathcal{M}}_\rho^+ \in \{0,1\}^{|L| \times |B_S| \times |\mathcal{E}_S|}$ is a third-order tensor whose element $\widetilde{\mathcal{M}}_\rho^+(i, y, \psi) = 1$ when the system capability $\epsilon_\psi \in \mathcal{E}_S$ injects operand $l_i \in L$ into buffer $b_{s_y} \in B_S$. ∎

In the context of the AMES, the operands are flows of matter and energy like coal, oil, natural gas, and electricity. The buffers are the point facilities like electric power plants and refineries. The capabilities are "subject + verb + object" sentences such as "NG refinery refines raw natural gas". The HFITs are important because they include all of the information necessary to produce: (1) a Formal Graph (FG) adjacency matrix (A_{B_S}) where point-facilities are connected via edge-facilities, and (2) a Hetero-Functional Graph (HFG) adjacency matrix A_ρ where the system's capabilities follow one another.

Definition 11 (*The Formal Graph Adjacency Matrix A_{B_S}*). The formal graph adjacency matrix $A_{B_S} \in \{0,1\}^{|B_S| \times |B_S|}$ is a matrix whose element $A_{B_S}(y_1, y_2) = 1$ when there is a physical connection between buffer y_1 and buffer y_2. ∎

Definition 12 (*Hetero-functional Adjacency Matrix [78]*). A square binary matrix A_ρ of size $|\mathcal{E}_S| \times |\mathcal{E}_S|$ whose element $A_\rho(\psi_1, \psi_2) \in \{0,1\}$ is equal to one when string $z_{\psi_1, \psi_2} \in \mathcal{Z}$ is available and exists. ∎

Each of these adjacency matrices is then calculated from the incidence tensors. The formal graph adjacency matrix A_{B_S} requires two steps. First, the two HFITs are summed along the operand dimension to produce two incidence matrices [81,82]:

$$M_B^+(y, \psi) = \sum_i^{|L|} \mathcal{M}_\rho^+(i, y, \psi) \tag{1}$$

$$M_B^-(y, \psi) = \sum_i^{|L|} \mathcal{M}_\rho^-(i, y, \psi) \tag{2}$$

It is important to recognize that the operand heterogeneity information is lost in this process. Then, these incidence matrices are multiplied.

This multiplication also results in a loss of information where the allocation of function onto form in the form of capabilities is dropped to describe the physical connection of buffers [81,82].

$$A_{B_S} = M_B^+ M_B^{-T} \tag{3}$$

In the meantime, the hetero-functional graph adjacency matrix A_ρ is calculated without loss of information after the HFITs have been matricized (or flattened) into hetero-functional incidence matrices M_ρ^+ and M_ρ^- with dimension $|L||B_S| \times |\mathcal{E}_S|$ [81,82].

$$A_\rho = M_\rho^{+T} M_\rho^- \tag{4}$$

While the FG adjacency matrix shows the physical connections from one point-resource (i.e. buffer) to another, the HFG adjacency matrix shows the logical sequence of capabilities one after the other. This results in the HFG adjacency matrix not just describing the physical connections but also the flow of functionality on matter and energy. This allows the model to describe the system on an asset-level basis defining the individual physical resources and the functionalities the resource performs. The use of HFGs in this way affords the modeler the ability to track the delivery of services as well as structural resilience when subject to formal and functional changes. As previous works have shown, the HFG allows for more comprehensive resilience analyses; be it for small electric power distribution systems [51] or for full-scale analysis of the American electric power system [52]. The open-source HFGT toolbox [80] provides an automated means for processing the input XML files at the formidable scale presented by the AMES.

4. Results: structural model statistics

Once created, the hetero-functional incidence tensor, the formal graph, and the hetero-functional graph for the four regions of NY, CA, TX, and the full USA can be compared. The following section presents these comparisons in the following subsections:

(1) Computational Performance of the HFGT Toolbox
(2) Formal Graph Statistics
(3) Hetero-Functional Graph Statistics
(4) Formal Graph Resource Distribution
(5) Hetero-Functional Capability Distribution
(6) Hetero-Functional Graph Process and Operand Degree Distribution

4.1. Computational performance of HFGT toolbox

The hetero-functional graph theory toolbox [51] is used as a computational tool that supports hetero-functional graph theory computations [68,74,81]. For this work, the exceptional scale of the AMES required extensive computational improvements to the HFGT toolbox so as to reduce computation times and memory used to produce the XML files and HFGs. For this work, all the models and data processing to produce the XML files was done on an iMac desktop with a 4 GHz Quad-Core Intel Core i7 processor and 32 GBs of RAM. The computation times of major code milestones are presented in Table 1. The resulting file sizes are also presented in Table 1. Note that the overwhelming majority of the computation is devoted to converting the GIS

Table 2

Formal graph statistics for the multi-modal energy systems in NY, CA, TX, and the USA.

	New York	California	Texas	United States
# of Operands	13	13	14	14
# of Buffers	7686	16,754	197,108	473,321
# of Edges	9115	20,674	179,895	511,802
Formal graph sparsity	1.5438e−4	7.3661e−5	4.6306e−6	2.2848e−6
Population (millions)	19.8	39.2	29.5	306.7
Land area (sqr miles)	54,556	163,696	268,597	3,119,884
Population density (ppl/sqr mile)	362.9	239.5	109.8	98.3
Buffers/land density (buffers/sqr mile)	0.1409	0.1023	0.7338	0.1518
Buffers/population density (buffers/ppl)	3.8818e−4	4.2740e−4	6.6816e−3	1.5432e−3
Edges/land density (edges/sqr mile)	0.1671	0.1263	0.6698	0.1641
Edges/population density (edges/ppl)	4.6035e−4	5.2740e−4	6.0981e−3	1.6687e−3

Table 3

Hetero-functional incidence tensor statistics for the multi-modal energy systems in NY, CA, TX, and the USA.

	New York	California	Texas	United States
# of Operands	13	13	14	14
# of Buffers	7686	16,754	19,7108	473,321
# of Capabilities	43,766	100,349	1,430,588	3,130,235
# of Elements in M_ρ^+.	40,921	94,377	1,422,463	3,028,855
# of Elements in M_ρ^-.	42,778	99,796	1,425,699	3,070,083
HFIT sparsity	4.6484e−5	2.2221e−5	1.4848e−6	6.7354e−7
HFAM sparsity	1.0618e−4	4.8221e−5	2.8641e−6	1.4259e−6
Capability land density (capabilities/sqr mile)	0.8022	0.6130	5.3262	1.0036
Capability population density (capabilities/ppl)	2.2104e−3	2.5599e−3	4.8495e−2	1.0206e−2

shape files into XML files. More specifically, automatic data cleaning and processing activities dominated the computation. In comparison,

Interestingly, the majority of the computation time came from processing and cleaning the GIS data. Specifically, the processing and cleaning post assigning clusters to points takes the largest amount of time when producing the XML. The actual assignment of clusters takes about $1/10th$ of the total time to produce the XMLs for all regions. When looking at the computation time of the HFGT Toolbox nearly all of the required time is spent calculating A_ρ. This calculation is ported over to Julia from Python through the use of CSV files, A_ρ is calculated, then ported back to Python via another CSV. The hetero-functional adjacency matrix is at the heart of the HFG as it represents all flows of capabilities with the system. In all, the HFGT toolbox computations themselves are highly optimized and are produced in under two hours on a moderately sized computer even for the full AMES dataset.

In line with the scale of each region, the smaller the region is, the faster both the XML file and HFG is produced; with NY being the fastest followed by CA, TX, and then the full USA. While CA is about twice the size of NY it took 4 times as long to produce the XML than NY. When comparing TX to NY it can be seen that TX is 26 times as large and takes 60 times as long to produce an XML. This suggests the computation efficiency of creating an XML is that of $2N$. When comparing the computation times for the HFGT Toolbox, the same trend emerges suggesting the toolbox also has a computational efficiency of $2N$. Additionally, seeing that TX takes about a quarter of the time to complete the HFGT toolbox as the USA, it suggests that TX composes about half of the American energy infrastructure. The significant portion of energy resources that TX commands is confirmed by the later sections.

4.2. Formal graph statistics

From the assessment of the computational intensity, the basic statistics of the formal graph models for the four regions can be compared relative to the size of their populations and land areas (Table 2).

As expected by the relative size of population and land area, Texas's energy infrastructure is the largest of the three states with California presenting the second largest and New York being the smallest in terms of the number of point-resources (buffers). The same ordering holds true for the number of edges connecting the buffers in each of the

states with TX presenting the most edges and NY presenting the least. All three states are of course included in the full USA model including all of their buffers and edges. There are 9115 edges in NY, 20,674 edges in CA, 179,895 edges in TX, and 511,802 edges in the USA. This corresponds to an adjacency matrix sparsity of 1.5438e−4, 7.3661e−5, 4.6306e−6, and 2.2848e−6 for NY, CA, TX, and the USA respectively. As expected, the larger the system the more sparse the adjacency matrix becomes. While the three states display energy infrastructures of differing scales, they all show a diverse energy mixture that utilizes the same 13 operands with Texas and the USA additionally utilizing solid biomass feedstock. Also, despite these absolute measures, NY's population density is significantly higher than that of Texas, California, or the USA with TX and the USA having the lowest population densities. This means that, if all else is held equal, New Yorkers receive more energy infrastructure benefits whereas Californians and Texans and the USA must spend more on energy infrastructure costs.

Table 2 also presents the buffers per land and population densities. When looking at the buffers per square mile density, CA had the smallest followed by NY, USA, and then TX reporting 0.1023, 0.1409, 0.1518, and 0.7338 respectively. This interesting result shows the impact that TX has on influencing the USA's energy infrastructure as a whole. TX being such a large percentage of the USA's infrastructure, it brings up the buffer per square mile density above other states like NY or CA. This is also seen in the buffers per population density with NY having the lowest followed by CA, USA, and TX respectively. These trends can also be tracked through the edges per square mile or population density. Just as with buffers, CA has the lowest edges per square mile density followed by NY, USA, and TX respectively. Then when comparing the edges per population NY and CA switch, presenting NY with the lowest density followed by CA, USA, and TX respectfully. These density measures can be used as a course indicator for the energy investment costs in each region. The lower the resource per population density, the greater the proxy cost per individual. Similarly, the lower the resource per square mile density, the more spread out the cost is across the region's landscape.

4.3. Hetero-functional graph statistics

The statistics of the hetero-functional graphs for NY, CA, TX, and the USA are presented in Table 3 similarly to those of the formal
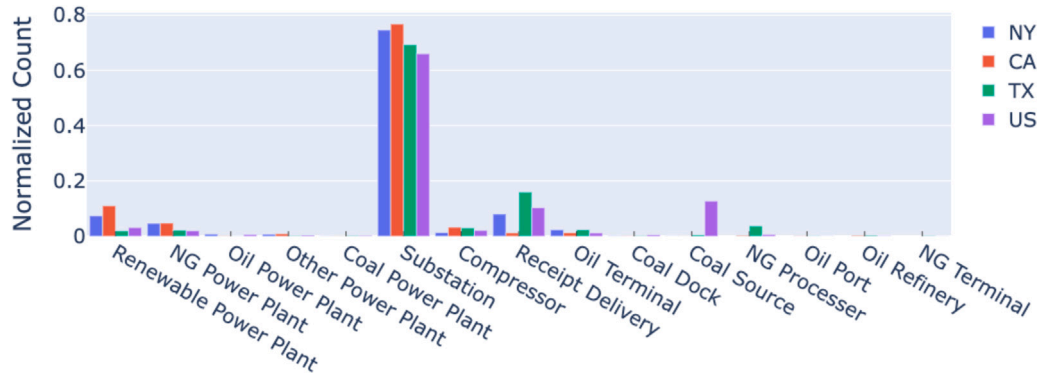
Fig. 3. Normalized distribution of buffer types for NY, CA, TX, and the full USA.

graphs. Trends of sparsity in the HFITs match those of the formal graph adjacency matrices. When comparing the number of capabilities in each model, NY has the least, followed by CA, TX, and the USA with $(43, 766)$, $(100, 349)$, $(1, 430, 588)$, and $(3, 130, 235)$ capabilities respectively. Unsurprisingly, the trend in the sparsity of the HFITs follows that of the formal adjacency matrix. NY's HFIT is twice as dense as CA's, which is 31 times as dense as TX's, and which is 69 times as dense as the USA's HFIT. Additionally, across all regions, the negative HFIT has more filled elements than the positive indicating a greater amount of capabilities that pull operands from capabilities rather than injecting them. A network topology with more capabilities that pull operands than inject operands is synonymous with a system that gathers and refines operands then distributes and decomposes them. In the largest sense, this mathematical result is consistent with our understanding of the AMES which collects raw energy commodities (especially coal, oil, and natural gas) and delivers them as relatively few "high-grade" energy commodities like electricity. The hetero-functional adjacency matrix can be calculated directly from the HFITs so as to link capabilities together. Again, as expected, these sparse matrices follow the same density trend as the HFIT with NY having the most dense matrix followed by CA, TX, and the USA with sparsity values of $1.0618e{-}4$, $4.8221e{-}5$, $2.8641e{-}6$, and $1.4259e{-}6$ respectfully.

Table 3 also presents the capabilities per land and population densities. When looking at the capabilities per square mile density, CA had the smallest followed by NY, USA, then TX reporting 0.6130, 0.8022, 1.0036, and 5.3262 respectively. The ordering of capabilities per square mile density is the same trend as presented with buffers per square mile in the formal graph statistics. Similarly, it shows the impact that states like TX have on influencing the full USA's energy infrastructure trends. Additionally, due to the nature of allocating function onto form, there are many more capabilities than buffers which brings the value of the density over land statistics up. The increase in density values is also seen in the capabilities per population density with NY having the lowest followed by CA, USA, and TX respectively. These density measures can be used as an indicator for the benefit of each region's energy system. The higher the capabilities per population density, the greater the amount of functionality (or benefits) from the infrastructure the population experiences. Similarly, the higher the capability per square mile density, the more infrastructure benefits can be accessed across the region's landscape.

4.4. Formal graph resource distribution

While it is important to assess the number of buffers (e.g. point energy facilities) in the multi-energy infrastructure of the three American states, it is also important to differentiate them by type. Fig. 3 shows that 74.6%, 76.8%, 69.3%, and 66% of the buffers in the formal graph are electric power substations in the states of NY, CA, TX, and the USA respectively. The high percentage of substations reflects the highly ubiquitous nature of the electric power system in all four regions. Furthermore, another 13.5%, 16.8%, 4.6%, and 6.4% of buffers are devoted towards electric power generation facilities (of various types) for NY, CA, TX, and the USA respectively. Because coal, oil, and natural gas are very dense approximate forms of energy their processing facilities for these types of energy have very strong economies of scale. Therefore, there is a trend towards centralization and a small number of point facilities for energy conversion. California, notably, has a greater shift towards the electric grid with a greater presence of substations and power plants than NY and TX. Meanwhile, Texas is notably a large producer and trader of fossil fuels in the USA and thus has the infrastructure to match. That NY has a greater reliance on oil and gas facilities is likely a byproduct of it being located in a colder climate. In opposition to the need for fossil fuels for space heating in cold environments, both TX and CA depend more on electricity for cooling spaces in their warmer climates and can make greater use of renewable energy resources for such a cause.

The first massive peak around substations reflects that there are more substations than any other type of node by far. In addition to identifying the prominent dependence on electric power in American life, the peak also shows that the transmission system comes a lot closer to the grid periphery of electric power systems than the corresponding distribution systems for coal, oil, or natural gas. By taking a data-driven approach to the Platts data, the discrepancies of how close each system's transmission level assets reach towards the system periphery becomes much more apparent. For the electric grid, the system boundary ends at substations; ignoring the distribution system assets and consumption and generation at homes and businesses. When looking at oil and natural gas, the system boundary stops at the terminals and ports where they are ultimately distributed outward. By taking a data-driven approach to the Platts data, tankers and smaller gas lines delivering fuel for residential and industrial use are also ultimately omitted. As a result, gas stations, homes, and other retail aspects of the oil and gas industries are not included. Additionally, with coal being sold between commercial entities and not going out to individual consumers, the number of coal facilities needed to distribute to the demand is quite limited. Despite the system boundary limitations of the Platts data, the analysis speaks to the state of the existing infrastructure and how much easier the electric power mode is to distribute than any other. The ease of electric power transportation, along with the potential reduction in carbon emissions, is one of the advantages of electrifying energy demands as a part of the sustainable energy transition.

Beyond the number and type of point-energy facilities, the formal graph also measures their interconnectedness. Despite the heterogeneity of point-energy facilities and the sparsity of the four formal graphs of each state and the full USA, Fig. 5 shows that the formal graph degree distributions for all four regions are remarkably similar. In contrast to the well-known power-law degree distribution for electric power
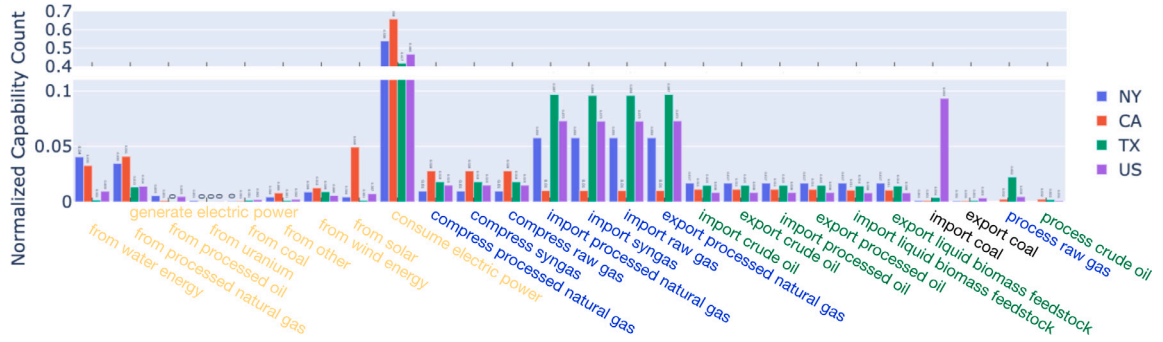
Fig. 4. Normalized capability distribution of NY, CA, TX, and the USA.
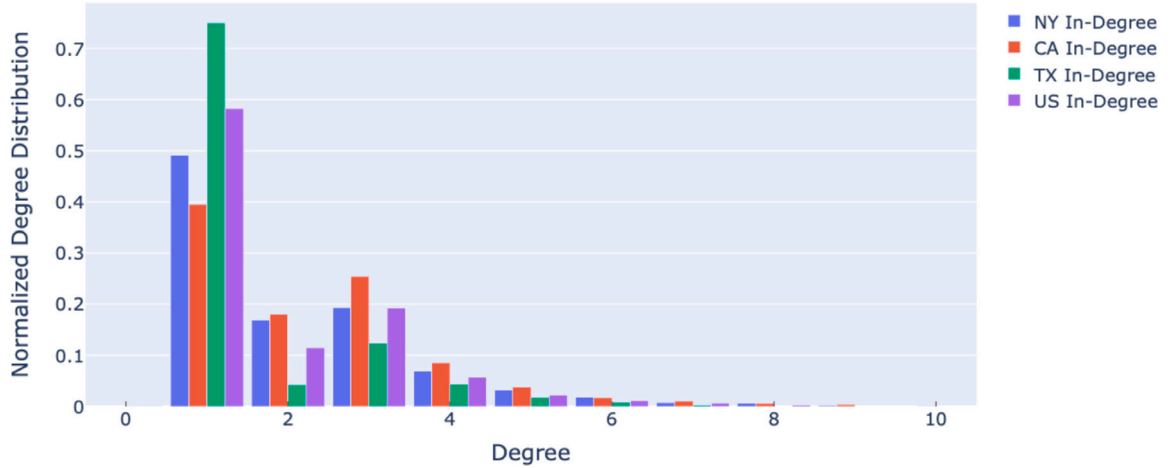

Fig. 5. Normalized formal graph degree distribution of NY, CA, TX, and the USA.

systems [6,83–86], these multi-energy systems have degree distributions with a rather unusual shape. Notability, all four regions peak at a single degree, then have a major lack of nodes with a degree 2. After the initial dip, the degree distribution jumps back up to nodes more prevalent with a degree 3 before tailing off into an exponential decay. Each energy subsystem likely contributes its own power-law degree distribution so that the degree distributions depicted in Fig. 5 are actually a composition of phenomena associated with the delivery of each energy commodity. These differing contributing factors are further investigated through the hetero-functional graph distributions reported in the following sections.

4.5. Hetero-functional capability distributions

In addition to looking at the type of nodes present in each region, the presence of different capabilities can also be compared. As such, the normalized counts of the different capabilities are presented in Fig. 4 for NY, CA, TX, and the full USA. Fig. 4 visualizes the prevalence of differing energy mixtures for each region as determined by their existing capabilities. Fig. 4 shows a large spike in electric power capabilities similar to the formal graph node distributions shown in Fig. 3. Another notable trend is that TX and the full USA have very similar capability distributions; presumably by virtue of the large size of the TX energy system relative to the full USA. The striking difference between TX and the USA is that the USA has an abundance of coal resources while TX has a very minimal amount. The lack of coal in TX is likely a result of the coal sources residing most prominently in the Appalachian region. Additionally, NY, TX, and the USA have a heavy reliance on natural gas functionality. All three regions have heavy spikes in functionality related to importing and exporting gas resources. NY, as a relatively cold northeastern state, relies heavily

on natural gas for space heating during the cold winters. In contrast, TX's natural gas functionality comes as a result of its large oil and gas economy. The USA's capability distribution mixes these two dependencies. Alternatively, CA is much more reliant on electric power generation as it has space conditioning requirements are largely driven by electrified cooling rather than natural-gas-based heating. Notably, the functionality of solar, hydro, and natural gas electric generation are most prominent as a result of CA's sustainable energy transition and the flexibility of natural gas electric generation to balance variable energy resources (VERs). These natural gas electric generation capabilities provide valuable ramping functionality to balance the electric grid through sharp upward and downward ramps in response to VER generation and in demand levels. These results are further confirmed and supported by the results in Fig. 6 which presents the normalized capacities of each electric generation capability normalized by region.

As the sustainable energy transition requires an electrification of the demands placed on the AMES, it becomes important to investigate the electric power generation mix. In accordance with single energy mode analysis, when the electric power generation capabilities in Fig. 4 are weighted by generation capacity, it produces Fig. 6 which shows the electric generation mix for the three states and the USA. The main source of generation is found to be natural gas for all four presented regions. The high presence of natural gas generation indicates that the need for quick response generation sources to support VERs for NY, CA, TX, and the USA in addition to low fuel costs has led to a large dependency on natural gas for electric power generation. Additionally, NY sees a larger generation capacity from processed oil and nuclear power facilities than the other regions. Again, the reliance on oil-fired generation stems from the need for "peaker" units that support electrified heating in the coldest winter periods. CA's commitment to renewable energy generation is demonstrated in its solar and hydro
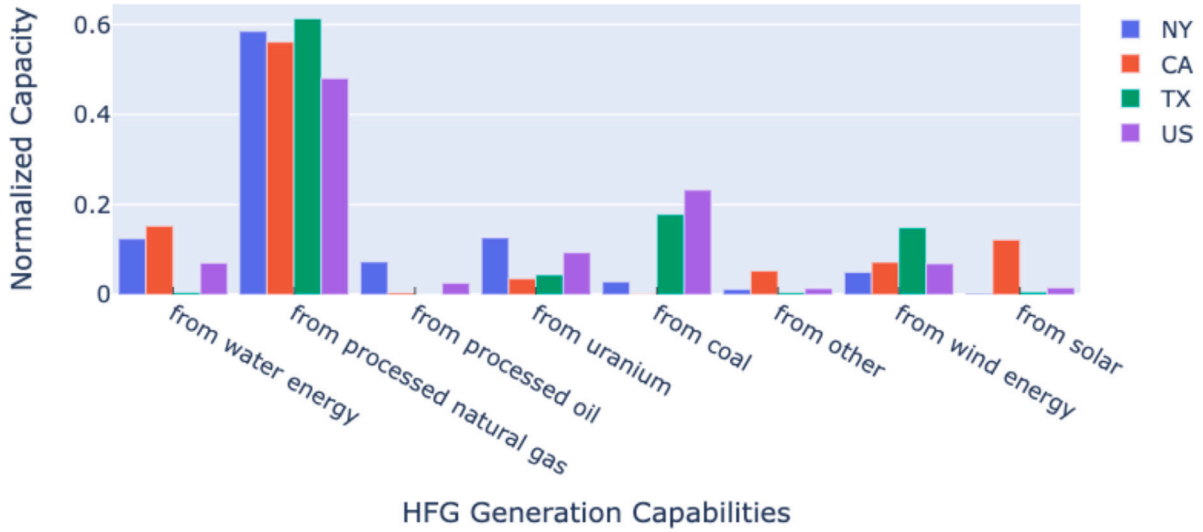
Fig. 6. Normalized electric generation capacity by fuel source for NY, CA, TX, and the full USA.

capacity. While Texas has a large investment in wind power. Both TX and the USA also have significant capacity in coal-fired electric power generation. In contrast, Fig. 6 shows CA's and NY's energy transition to cleaner fuel sources and away from coal. Finally, the electric power generation mix in Fig. 6 demonstrates that when HFGT is restricted to a single commodity network (e.g. electricity), it is able to reproduce the familiar analytics associated with individual (and often siloed) infrastructures.

Just as the degree distributions of the AMES FGs can be presented so can the degree distributions for the HFGs. Figs. 7(a) and 7(b) show the AMES' HFG in-degree and out-degree distributions respectively. As HFGs describe the logical sequence of capabilities, they are fundamentally directed in nature. As such, both the in-degree and out-degree must be presented. These plots present similar trends to those seen in the FG degree distributions shown in Fig. 5. Neither the FG degree distribution nor the HFG degree distributions follow a traditional power law decay, but the long tails in each plot suggest that there are some underlying power-law behaviors. From the HFG degree distributions, it is clear that there is some combination of exponentially decaying degree distributions from each of the four subsystems composing the AMES. Interestingly, the general shape of both the in and out-degree plots are very similar. Both plots have normalized distributions that bottom out by degree 20 and have peaks at degrees 2 and 4. Additionally, all four regions also have a major dip in their degree distributions at degree 3. While the four regions all seem to have some underlying power law behaviors, each region presents its own combination of decays. The differing degree distribution combinations suggest that there are geographical dependencies that influence the degree distributions. Ultimately, while the results in Fig. 7 confirm those of Fig. 5, the HFG model invites further investigation (in the following section) by classifying the HFG's capabilities by their underlying process.

4.6. Hetero-functional graph process degree distribution

Figs. 8 and 9 present three-dimensional (3D) degree distributions of the AMES' HFG where the third dimension classifies the node capabilities by their underlying process. This choice of classification is of critical analytical importance. As discussed extensively in prior works [68,81], capabilities can be associated with either point-type buffers or line-type transportation resources. Therefore, a classification scheme based on buffers is both incomplete and logically inconsistent with the nature of a HFG. In the meantime, a classification scheme based on the capabilities' (input or output) operands runs the risk of double counting the capabilities because each capability can have more

than one type of operand (e.g. a natural gas electric power plant). Therefore, the most straightforward way of classifying capabilities into mutually exclusive and totally exhaustive sets while preserving the objectivity of the underlying statistics is based upon the capabilities' underlying process.

Returning to Figs. 8 and 9, the former presents the 3D in-degree distribution while the latter presents the 3D out-degree distribution. Both figures reveal, for the first time, the underlying nature of the AMES' hetero-functional structure. With rare exception, the AMES exhibits a power-law degree distribution for each set of capabilities classified by process. These figures provide a fascinating empirical result. Previous works demonstrate that HFGs of single operand networks have power-law degree distributions just like formal graphs of single operand networks [51,52]. Now, this paper shows that a HFG of a multi-operand network exhibits a power law degree distribution for each type of process so that the final HFG degree distribution in Fig. 7 shows a superposition of the degree distributions associated with each type of process. In contrast, such an empirical result cannot be reached using a formal graph with buffer-nodes because of the information loss caused by the sums in Eqs. (1) and (2). In other words, HFGT has confirmed well-known results in the network science literature and successfully generalized them for systems that are fundamentally hetero-functional in nature.

This theoretical insight has direct practical relevance to the fundamental systems science principles underlying the sustainable energy transition. The 3D degree distributions in Figs. 8 and 9 show that the AMES' structure is evolving at different rates depending on the underlying process of each capability. In other words, from a graph theory perspective, the sustainable energy transition can be understood as *network phenomena* where the decommissioning of carbon-intensive processes (e.g. oil and gas refining and coal-fired electric power generation) is occurring at a certain rate, and carbon-light processes (e.g. wind and solar power generation) is occurring at a different rate. Similarly, broad trends towards electrification (or fuel switching) is a superposition of the network phenomena driving the consumption of carbon-intensive fuels relative to consuming electricity. Said differently, HFGT points to the network phenomena underpinning "infrastructure lock-in" like homeowners who wish to switch to heat-pump technology but find themselves "locked in" to furnace-based heating. Similarly, HFGT points to the network phenomena underpinning "infrastructure platforms" like car owners who wish to buy an electric vehicle but find themselves worrying about local charging infrastructure. Ultimately, a deeper study of the power laws in Figs. 8 and 9 is warranted because the relative rates underpin the success of the sustainable energy
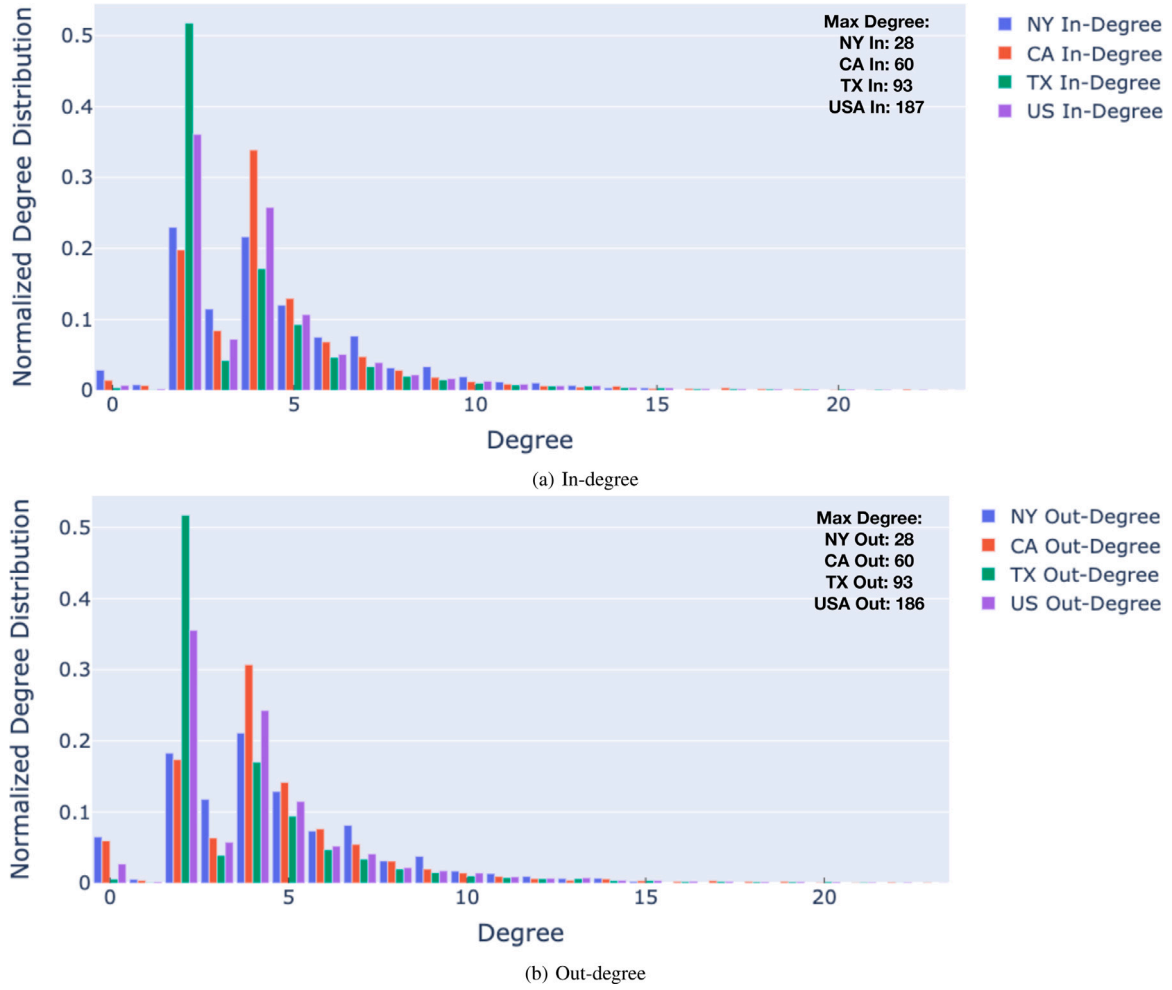
(a) In-degree

(b) Out-degree

Fig. 7. Normalized hetero-functional graph degree distribution of NY, CA, TX, and the USA.

transition. From a policy perspective, the sustainable energy transition requires that the rate of integrating certain types of sustainable energy capabilities be "sped up" relative to more carbon-intensive ones.

5. Conclusions and policy implications

This paper uses a data-driven, MBSE-guided approach to develop open-source software that produces open structural models of the American Multi-modal Energy System. It is part of a larger NSF project entitled "American Multi-Modal Energy System Synthetic & Simulated Data (AMES-3D)" which seeks to produce open-source structural and behavioral models of the American Multi-modal Energy System. The creation of open-source software and open-data models of the AMES fills an important need in open, citizen-based science in America's sustainable energy transition. It also provides one of the few multi-energy system datasets in which to advance fundamental methods. The AMES' structural models are inferred from the Platts Map Data Pro GIS dataset and is complemented by the previously developed American Multi-modal Energy System Reference Architecture [46]. Together, these two data sources serve as the basis for an XML-based input data file for the open-source hetero-functional graph theory toolbox.

This paper specifically reports the hetero-functional incidence tensor, the formal graph adjacency matrix, and hetero-functional graph adjacency matrix statistics for the multi-energy infrastructure systems for the states of NY, CA, TX, and the full USA. Here, the application of hetero-functional graph theory facilitates a nuanced analysis that respects the heterogeneity in this highly interdependent system-of-systems. The paper finds that the geography and sustainable energy

policies of the states are deeply reflected in the structure of their multi-energy infrastructure. Because New York's cold northeastern climate drives heating demand, it has a multi-energy system with a greater emphasis on oil and gas. In the meantime, California's warm climate is reflected in a multi-energy system with a greater emphasis on electric power systems. Additionally, Texas has a large oil and gas economy and thus has a large percentage of energy infrastructure pertaining to these fossil fuels but also has built out of a tremendous amount of wind energy for electric generation. Along these lines, California and Texas have also geared their natural gas resources infrastructure towards electric power generation to support their growing reliance on variable energy resources. These trends appear as components of the AMES as a whole. Additionally, states with a large energy infrastructure like Texas have a greater impact on shaping the USA's energy infrastructure. It is also important to note the import and export functionality of these fuel sources are also very prominent. Identifying the abundance of import and export functionality is important for the evolution of a multi-modal energy system for two reasons. Import and export functionality provides open interfaces to new modes of energy delivery such as hydrogen. Important export functionality in a HFG also creates a theoretical foundation from which to investigate energy systems with relatively large import/export economies (e.g. Australia). Finally, through the analysis of HFG degree distributions and their associated processes, it becomes clear that power-law degree distributions are fundamentally tied to the processes rather than point-resources (i.e. buffers).

From the perspective of understanding the dynamics of the sustainable energy transition, the power laws imply that for each type of
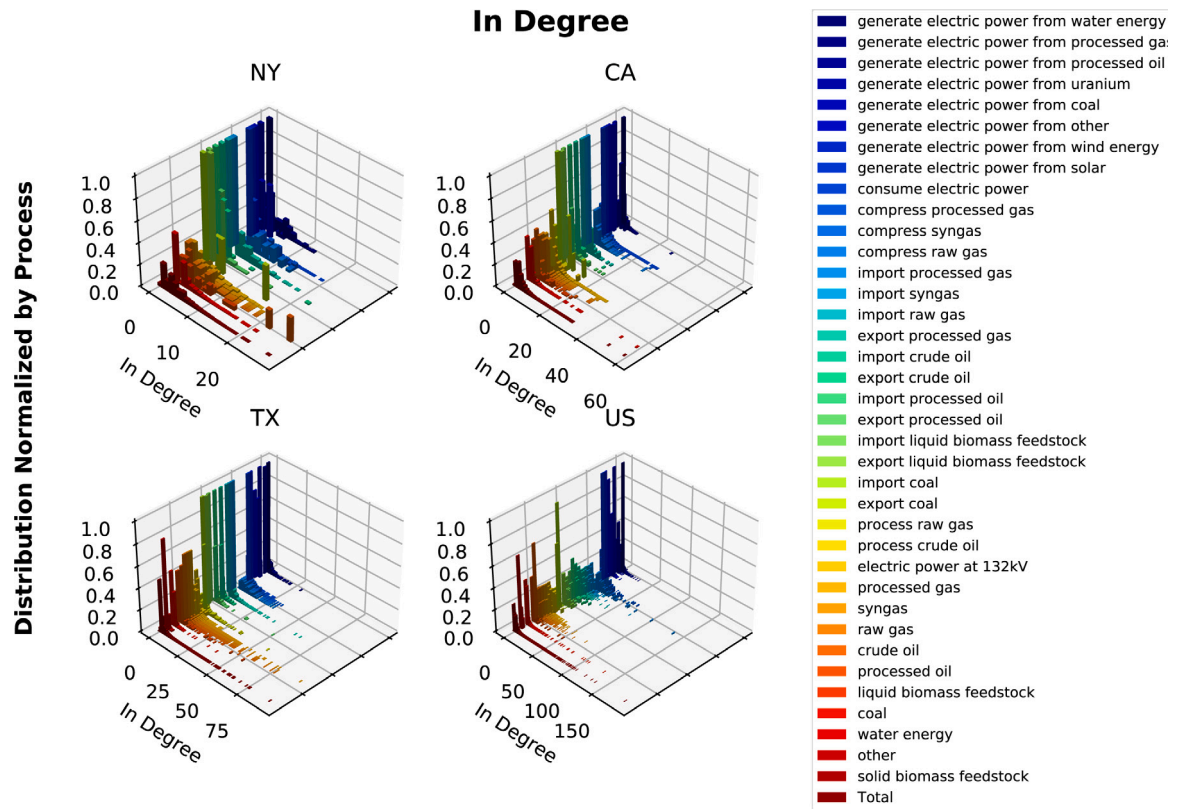
In Degree

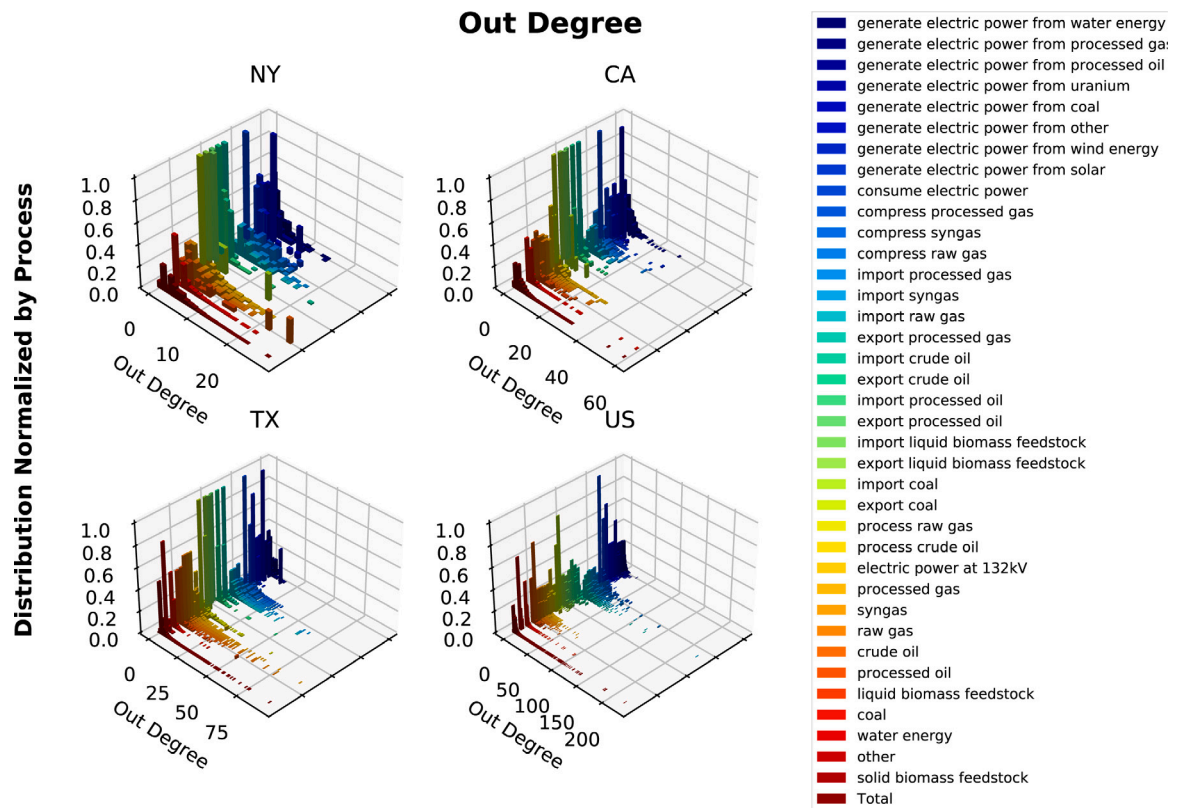**Fig. 8.** In-Degree distribution of all Processes.

Out Degree

Fig. 9. Out-Degree distribution of all Processes.

process, the "popular" (e.g. most well-connected) become even more so. The degree distribution power law for each process reinforces the connections of the capabilities with more connected capabilities. The AMES' HFG has a power law associated with every type of process and each power law reflects a rate of evolution or adoption.

Consequently, many new sustainable energy process technologies may require policies, at least initially, that support them relative to incumbent process technologies. These initial policies are likely to instigate network-driven positive feedback loops that accelerate adoption rates. For example, as renewable energy sources are adopted, they expand the electric grid itself, which in turn supports electrification on the AMES' demand side. These implications are also applicable to hydrogen as it becomes a more viable and prominent technology. As this work was driven by the Platts GIS data, it was ultimately confined to incumbent conventional energy sources, rather than new energy pathways tied to hydrogen, synthetic fuels, and bio-energy. Naturally, the transparent nature of the AMES reference architecture lends itself to revision so as to include these new energy carriers based upon first-principle process physics. Finally, the power law results found in Figs. 8 and 9 present a ripe opportunity for further investigation.

Beyond the structural analysis presented here, the AMES' structural model presents multiple avenues for future open-science research. First, expanding on previous works examining the resilience of electric power systems [51,52], this hetero-functional graph model can be used to examine the resilience of the entire AMES to cascading failures and attacks on multiple modes of energy. Additionally, behavioral data can be incorporated so as to develop a physically-informed machine-learning behavioral model of the AMES [87] that rigorously studies the AMES' sustainability and resilience. Finally, scenarios varying the energy mixtures seen in these regions and across the entire USA can be investigated to explore potential sustainable energy transitions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

Appendix A

See Tables 4 and 5.

Table 4
This table presents the normalization factors for Figs. 3–7 for the regions NY, CA, TX, and the USA.

	New York	California	Texas	United States
Normalization factor Fig. 3	3.15e−4	1.37e−4	1.14e−4	7.93e−6
Normalization factor Fig. 4	1.30e−4	5.97e−5	5.07e−6	2.11e−6
Normalization factor Fig. 5	2.27e−4	1.17e−4	6.88e−5	5.60e−6
Normalization factor Fig. 6	2.71e−5	1.52e−5	8.59e−6	9.25e−7
Normalization factor Fig. 7A	2.29e−5	9.96e−6	6.99e−7	3.24e−7
Normalization factor Fig. 7B	2.29e−5	9.96e−6	6.99e−7	3.24e−7

Table 5
This table presents the normalization factors for Figs. 8 and 9 for the regions NY, CA, TX, and the USA.

	NY In	NY Out	CA In	CA Out	TX In	TX Out	USA In	USA Out
Generate electric power from water energy	5.62e−3	5.62e−3	3.60e−3	3.60e−3	4.76e−2	4.76e−2	7.70E−4	7.70E−4
Generate electric power from processed gas	6.17e−3	6.17e−3	2.83e−3	2.83e−3	5.05e−3	5.05e−3	5.20E−4	5.20E−4
Generate electric power from processed oil	4.17e−2	4.17e−2	0.125	0.125	0.33	0.33	1.63E−3	1.63E−3
Generate electric power from uranium	0.25	0.25	1	1	0.5	0.5	1.96E−2	1.96E−2
Generate electric power from coal	0.5	0.5	0.2	0.2	5.26e−2	5.26e−2	3.31E−3	3.31E−3
Generate electric power from other	5.26e−2	5.26e−2	1.47e−2	1.47e−2	5.88e−2	5.88e−2	2.97E−3	2.97E−3
Generate electric power from wind energy	2.56e−2	2.56e−2	9.43e−3	9.43e−3	7.69e−3	7.69e−3	2.10E−3	2.10E−3
Generate electric power from solar	5.26e−2	5.26e−2	2.39e−3	2.39e−3	5.56e−2	5.56e−2	1.56E−3	1.56E−3
Consume electric power	4.23e−4	4.23e−4	1.79e−4	1.79e−4	1.65e−4	1.65e−4	1.56E−4	1.56E−4
Compress processed gas	2.38e−2	2.38e−2	4.22e−3	4.22e−3	3.81e−3	3.81e−3	1.19E−2	1.19E−2
Compress syngas	2.38e−2	2.38e−2	4.22e−3	4.22e−3	3.81e−3	3.81e−3	1.19E−2	1.19E−2
Compress raw gas	2.38e−2	2.38e−2	4.22e−3	4.22e−3	3.81e−3	3.81e−3	1.19E−2	1.19E−2
Import processed gas	3.94e−3	3.94e−3	1.16e−2	1.16e−2	7.11e−4	7.11e−4	2.45E−3	2.45E−3
Import syngas	3.94e−3	3.94e−3	1.16e−2	1.16e−2	7.11e−4	7.11e−4	2.45E−3	2.45E−3
Import raw gas	3.94e−3	3.94e−3	1.16e−2	1.16e−2	7.11e−4	7.11e−4	2.45E−3	2.45E−3
Export processed gas	3.94e−3	3.94e−3	1.16e−2	1.16e−2	7.11e−4	7.11e−4	2.45E−3	2.45E−3
Import crude oil	1.35e−2	1.35e−2	1.02e−2	1.02e−2	4.63e−3	4.63e−3	2.13E−2	2.13E−2
Export crude oil	1.35e−2	1.35e−2	1.02e−2	1.02e−2	4.63e−3	4.63e−3	2.13E−2	2.13E−2
Import processed oil	1.35e−2	1.35e−2	1.02e−2	1.02e−2	4.63e−3	4.63e−3	2.13E−2	2.13E−2
Export processed oil	1.35e−2	1.35e−2	1.02e−2	1.02e−2	4.63e−3	4.63e−3	2.17E−2	2.17E−2
Import liquid biomass feedstock	1.35e−2	1.35e−2	1.10e−2	1.10e−2	4.83e−3	4.83e−3	2.22E−2	2.22E−2
Export liquid biomass feedstock	1.35e−2	1.35e−2	1.10e−2	1.10e−2	4.83e−3	4.83e−3	2.22E−2	2.22E−2
Import coal	0.25	0.25	0.125	0.125	1.82e−2	1.82e−2	2.66E−3	2.66E−3
Export coal	0.33	0.33	0.17	0.17	6.67e−2	6.67e−2	0.05	0.05
Process raw gas	0	0	5.26e−2	5.26e−2	3.10e−3	3.10e−3	0.04	0.04
Process crude oil	0	0	5.26e−2	5.26e−2	3.85e−2	3.85e−2	0.2	0.2
Transport electric power	9.95e−5	9.95e−5	3.44e−5	3.44e−5	1.31e−4	1.31e−4	7.73E−5	7.73E−5
Transport processed gas	1.52e−4	1.52e−4	6.43e−5	6.43e−5	1.63e−5	1.63e−5	2.07E−4	2.07E−4
Transport syngas	1.65e−4	1.65e−4	6.77e−5	6.77e−5	1.63e−5	1.63e−5	2.09E−4	2.09E−4
Transport raw gas	1.65e−4	1.65e−4	6.77e−5	6.77e−5	1.63e−5	1.63e−5	2.07E−4	2.07E−4
Transport crude oil	8.47e−3	8.47e−3	4.82e−4	4.82e−4	2.73e−4	2.73e−4	2.11E−3	2.11E−3

Table 5 (continued).

	NY In	NY Out	CA In	CA Out	TX In	TX Out	USA In	USA Out
Transport processed oil	3.23e−3	3.23e−3	5.92e−4	5.92e−4	5.37e−3	5.37e−3	1.05E−2	1.05E−2
Transport liquid biomass feedstock	0	0	0	0	0	0	0	0
Transport coal	9.83e−5	9.83e−5	7.17e−5	7.17e−5	1.14e−3	1.14e−3	5.16e−4	5.16e−4
Transport water energy	0	0	0	0	0	0	0	0
Transport Other	0	0	0	0	0.01	0.01	4.76E−2	4.76E−2
Transport solid biomass feedstock	0	0	0	0	0	0	0	0
Total	2.28e−5	2.28e−5	9.97e−6	9.97e−6	6.99e−7	6.99e−7	3.24e−7	3.24e−7

References

[1] E.J. Moniz, A way forward on the climate crisis energy insecurity, 2022, [Online]. Available: https://www.bostonglobe.com/2022/07/18/opinion/way-forward-climate-crisis-energy-insecurity/.

[2] C. Action, Framework convention on climate change, 2010.

[3] WEC, World energy trilemma 2016: Defining measures to accelerate the energy transition, 2016.

[4] G. Bridge, S. Bouzarovski, M. Bradshaw, N. Eyre, Geographies of energy transition: Space, place and the low-carbon economy, Energy policy 53 (2013) 331–340.

[5] A. Vespignani, The fragility of interdependency, Nature 464 (7291) (2010) 984–985, [Online]. Available: http://dx.doi.org/10.1038/464984a.

[6] R. Albert, I. Albert, G.L. Nakarado, Structural vulnerability of the North American power grid, Phys. Rev. E 69 (2) (2004) 025103.

[7] G. Dong, R. Du, L. Tian, R. Liu, Robustness of network of networks with interdependent and interconnected links, Physica A 424 (2015) 11–18.

[8] P. Jean-Baptiste, R. Ducroux, Energy policy and climate change, Energy Policy 31 (2) (2003) 155–166.

[9] E. Kriegler, G. Luderer, N. Bauer, L. Baumstark, S. Fujimori, A. Popp, J. Rogelj, J. Strefler, D.P. Van Vuuren, Pathways limiting warming to 1.5c: a tale of turning around in no time? Phil. Trans. R. Soc. A 376 (2119) (2018) 20160457.

[10] R. Lempert, B.L. Preston, J. Edmonds, L. Clarke, T. Wild, M. Binsted, E. Diringer, B. Townsend, Pathways to 2050 Alternative Scenarios for Decarbonizing the U.S. Economy, Climate Innovation 2050, Center for Climate and Energy Solutions, 2019.

[11] J. Rogers, K. Averyt, S. Clemmer, M. Davis, F. Flores-Lopez, D. Kenney, J. Macknick, N. Madden, J. Meldrum, S. Sattler, E. Spanger-Siegfried, Water-Smart Power: Strengthening the U.S. Electricity System in a Warming World, Tech. Rep., Union for Concerned Scientists, Cambridge, MA, 2013, pp. 1–58.

[12] P. Lara, M. Sánchez, J. Villalobos, Enterprise modeling and operational technologies (OT) application in the oil and gas industry, J. Ind. Inf. Integr. 19 (2020) 100160, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2452414X20300352.

[13] M. Panteli, E.A.M. Ceseña, R. Moreno, P. Mancarella, Infrastructure planning under uncertainty: flexibility, resilience and multi-energy systems application tutorial, 2019.

[14] D. Mejia-Giraldo, J. Villarreal-Marimon, Y. Gu, Y. He, Z. Duan, L. Wang, Sustainability and resiliency measures for long-term investment planning in integrated energy and transportation infrastructures, J. Energy Eng. (USA), J. Energy Eng. 138 (2) (2012) 87–94, [Online]. Available: http://dx.doi.org/10.1061/(ASCE)EY.1943-7897.0000067.

[15] G.M. Masters, Renewable and Efficient Electric Power Systems, John Wiley & Sons, 2013.

[16] S. Mokhatab, W.A. Poe, Handbook of Natural Gas Transmission and Processing, Gulf Professional Publishing, 2012.

[17] M.V. Lurie, Modeling of Oil Product and Gas Pipeline Transportation, Wiley-VCH Verlag GmbH & Co. KGaA, 2009.

[18] EIA, Coal Market Module of the National Energy Modeling System: Model Documentation 2014, Independent Statistics and Analysis, U.S. Energy Information Administration, Department of Energy, 2014.

[19] E. Priyanka, S. Thangavel, X.-Z. Gao, N. Sivakumar, Digital twin for oil pipeline risk estimation using prognostic and machine learning techniques, J. Ind. Inf. Integr. (2021) 100272, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2452414X21000704.

[20] EIA, Annual Energy Outlook 2020, Independent Statistics and Analysis, U.S. Energy Information Administration, Department of Energy, 2020.

[21] EIA, Availability of the National Energy Modeling System (NEMS) Archive, Tech. Rep., United States Energy Information Administration, 2017, [Online]. Available: https://www.eia.gov/outlooks/aeo/info_nems_archive.cfm.

[22] A.M. Farid, W.N. Lubega, W. Hickman, Opportunities for energy-water nexus management in the Middle East and North Africa, Elementa 4 (134) (2016) 1–17, [Online]. Available: http://dx.doi.org/10.12952/journal.elementa.000134.

[23] W.N. Lubega, A.M. Farid, A reference system architecture for the energy-water nexus, IEEE Syst. J. PP (99) (2014) 1–11, [Online]. Available: http://dx.doi.org/10.1109/JSYST.2014.2302031.

[24] W.N. Lubega, A.M. Farid, Quantitative engineering systems model and analysis of the energy-water nexus, Appl. Energy 135 (1) (2014) 142–157, [Online]. Available: http://dx.doi.org/10.1016/j.apenergy.2014.07.101.

[25] J.R. Thompson, D. Frezza, B. Necioglu, M.L. Cohen, K. Hoffman, K. Rosfjord, Interdependent critical infrastructure model (ICIM): An agent-based model of power and water infrastructure, Int. J. Crit. Infrastruct. Prot. 24 (2019) 144–165.

[26] G.A. Munoz-Hernandez, S.P. Mansoor, D.I. Jones, Modelling and Controlling Hydropower Plants, Springer, London, 2013.

[27] W.N. Lubega, An Engineering Systems Approach to the Modeling and Analysis of the Energy-Water Nexus, (Master's Thesis), Masdar Institute of Science & Technology, 2014, pp. 1–11.

[28] W.N. Lubega, A. Santhosh, A.M. Farid, K. Youcef-Toumi, Opportunities for integrated energy and water management in the GCC – a keynote paper, in: EU-GCC Renewable Energy Policy Experts' Workshop, no. December, Masdar Institute, Abu Dhabi, UAE, 2013, pp. 1–33, [Online]. Available: http://www.grc.net/data/contents/uploads/Opportunities_for_Integrated_Energy_and_Water_Management_in_the_GCC_5874.pdf.

[29] W.N. Lubega, A. Santhosh, A.M. Farid, K. Youcef-Toumi, An integrated energy and water market for the supply side of the energy-water nexus in the engineered infrastructure, in: ASME 2014 Power Conference, Baltimore, MD, 2014, pp. 1–11, [Online]. Available: http://dx.doi.org/10.1115/POWER2014-32075.

[30] W.N. Lubega, A.M. Farid, An engineering systems sensitivity analysis model for holistic energy-water nexus planning, in: ASME 2014 Power Conference, Baltimore, MD, 2014, pp. 1–10, [Online]. Available: http://dx.doi.org/10.1115/POWER2014-32076.

[31] A.M. Farid, W.N. Lubega, Powering and watering agriculture: Application of energy-water nexus planning, in: GHTC 2013: IEEE Global Humanitarian Technology Conference, Silicon Valley, CA, USA, 2013, pp. 1–6, [Online]. Available: http://dx.doi.org.libproxy.mit.edu/10.1109/GHTC.2013.6713689.

[32] A. Santhosh, A.M. Farid, A. Adegbege, K. Youcef-Toumi, Simultaneous co-optimization for the economic dispatch of power and water networks, in: The 9th IET International Conference on Advances in Power System Control, Operation and Management, Hong Kong, China, 2012, pp. 1–6, [Online]. Available: http://dx.doi.org/10.1049/cp.2012.2148.

[33] A. Santhosh, A.M. Farid, K. Youcef-Toumi, The impact of storage facilities on the simultaneous economic dispatch of power and water networks limited by ramping rates, in: IEEE International Conference on Industrial Technology, Cape Town, South Africa, 2013, pp. 1–6, [Online]. Available: http://dx.doi.org/10.1109/ICIT.2013.6505794.

[34] A. Santhosh, A.M. Farid, K. Youcef-Toumi, The impact of storage facility capacity and ramping capabilities on the supply side of the energy-water nexus, Energy 66 (1) (2014) 1–10, [Online]. Available: http://dx.doi.org/10.1016/j.energy.2014.01.031.

[35] W. Hickman, A. Muzhikyan, A.M. Farid, The synergistic role of renewable energy integration into the unit commitment of the energy water nexus, Renew. Energy 108 (1) (2017) 220–229, [Online]. Available: http://dx.doi.org/10.1016/j.renene.2017.02.063.

[36] D. Murrant, A. Quinn, L. Chapman, The water-energy nexus: future water resource availability and its implications on UK thermal power generation, Water Environ. J. (2015).

[37] M. Wakeel, B. Chen, Energy consumption in urban water cycle, Energy Procedia 104 (2016) 123–128, [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1876610216315776.

[38] J. Macknick, R. Newmark, G. Heath, K.C. Hallett, Operational water consumption and withdrawal factors for electricity generating technologies: a review of existing literature, Environ. Res. Lett. 7 (4) (2012) 045802.

[39] J. Macknick, S. Sattler, K. Averyt, S. Clemmer, J. Rogers, The water implications of generating electricity: water use across the United States based on different electricity pathways through 2050, Environ. Res. Lett. 7 (4) (2012) 045803.

[40] K. Averyt, J. Fisher, A. Huber-Lee, A. Lewis, J. Macknick, N. Madden, J. Rogers, S. Tellinghuisen, Freshwater use by US Power Plants: Electricity's Thirst for a Precious Resource, Tech. Rep., Union of Concerned Scientists, Cambridge, MA, USA, 2011.

[41] V.C. Tidwell, J. Macknick, K. Zemlick, J. Sanchez, T. Woldeyesus, Transitioning to zero freshwater withdrawal in the U.S. for thermoelectric generation, Appl. Energy 131 (2014) 508–516, [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0306261913009215.

[42] S.A. Roosa, Sustainable Development Handbook, The Fairmont Press, Inc., 700 Indian Trail Lilbum, GA, 2008, pp. 1–497.

[43] J.K. Barry W. Johnson, Dear colleague letter: Simulated andsynthetic data for infrastructure modeling, 2017.

[44] D. Dori, Model-Based Systems Engineering with OPM and SysML, Springer, 2015.

[45] S. Friedenthal, A. Moore, R. Steiner, A Practical Guide To SysML: The Systems Modeling Language, second ed., Morgan Kaufmann, Burlington, MA, 2011, p. 640.

[46] D.J. Thompson, A.M. Farid, A reference architecture for the American multi-modal energy system, 2020, arXiv preprint arXiv:2012.14486.

[47] F. Plazas-Niño, N. Ortiz-Pimiento, E. Montes-Páez, National energy system optimization modelling for decarbonization pathways analysis: A systematic literature review, Renew. Sustain. Energy Rev. 162 (2022) 112406, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032122003148.

[48] A. Millot, A. Krook-Riekkola, N. Maïzi, Guiding the future energy transition to net-zero emissions: Lessons from exploring the differences between France and Sweden, Energy Policy 139 (2020) 111358, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301421520301154.

[49] A. Dedinec, V. Taseska-Gjorgievska, N. Markovska, T. Obradovic Grncarovska, N. Duic, J. Pop-Jordanov, R. Taleski, Towards post-2020 climate change regime: Analyses of various mitigation scenarios and contributions for Macedonia, Energy 94 (2016) 124–137, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S036054421501453X.

[50] D. Yangka, M. Diesendorf, Modeling the benefits of electric cooking in bhutan: A long term perspective, Renew. Sustain. Energy Rev. 59 (2016) 494–503, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032115016482.

[51] D.J. Thompson, W.C. Schoonenberg, A.M. Farid, A hetero-functional graph analysis of electric power system structural resilience, in: 2020 IEEE Power Energy Society Innovative Smart Grid Technologies Conference, ISGT, 2020, pp. 1–5.

[52] D.J. Thompson, W.C. Schoonenberg, A.M. Farid, A hetero-functional graph resilience analysis of the future American electric power system, IEEE Access 9 (2021) 68837–68848.

[53] Platts, Platts Energy Map Data Pro, Tech. Rep., S&P Global Platts, 2017, [Online]. Available: https://www.spglobal.com/platts/en/products-services/oil/map-data-pro.

[54] EIA, California State Energy Profile, EIA U.S. States Overview, 2021.

[55] EIA, New York State Energy Profile, EIA U.S. States Overview, 2021.

[56] EIA, Texas State Energy Profile, EIA U.S. States Overview, 2022.

[57] R. Cloutier, G. Muller, D. Verma, R. Nilchiani, E. Hole, M. Bone, The concept of reference architectures, Syst. Eng. 13 (1) (2010) 14–27.

[58] S. Oberle, R. Elsland, Are open access models able to assess today's energy scenarios? Energy Strategy Rev. 26 (2019) 100396, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2211467X19300896.

[59] P. Lopion, P. Markewitz, M. Robinius, D. Stolten, A review of current challenges and trends in energy systems modeling, Renew. Sustain. Energy Rev. 96 (2018) 156–166, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032118305537.

[60] M. Howells, H. Rogner, N. Strachan, C. Heaps, H. Huntington, S. Kypreos, A. Hughes, S. Silveira, J. DeCarolis, M. Bazillian, A. Roehrl, OSeMOSYS: The open source energy modeling system: An introduction to its ethos, structure and development, Energy Policy 39 (10) (2011) 5850–5870, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301421511004897. Sustainability of biofuels.

[61] EIA, The National Energy Modeling System: An Overview 2018, Independent Statistics and Analysis, U.S. Energy Information Administration, Department of Energy, 2019.

[62] E3MLab, PRIMES model version 2018 detailed model description, E3MLAB (2018).

[63] A. Muzhikyan, S.O. Muhanji, G. Moynihan, D. Thompson, Z. Berzolla, A.M. Farid, ISO New England System Operational Analysis and Renewable Energy Integration Study: Final Project Report, Tech. Rep., Laboratory for Intelligent Integrated Networks of Engineering Systems. Thayer School of Engineering at Dartmouth, Hanover, NH, USA, 2018, pp. 1–80, [Online]. Available: http://engineering.dartmouth.edu/liines/resources/TechReports/SPG-W07.pdf.

[64] S.O. Muhanji, C. Barrows, J. Macknick, A.M. Farid, An enterprise control assessment case study of the energy-water nexus for the ISO new England system, Renew. Sustain. Energy Rev. 141 (2021) 110766–110785, [Online]. Available: http://dx.doi.org/10.1016/j.rser.2021.110766.

[65] J.N.P. van Stralen, F. Dalla Longa, B.W. Daniëls, K.E.L. Smekens, B. van der Zwaan, OPERA: a new high-resolution energy system model for sector integration research, Environ. Model. Assess. 26 (6) (2021) 873–889, [Online]. Available: http://dx.doi.org/10.1007/s10666-020-09741-7.

[66] G. Limpens, S. Moret, H. Jeanmart, F. Maréchal, EnergyScope TD: A novel open-source model for regional energy systems, Appl. Energy 255 (2019) 113729, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261919314163.

[67] C.B. Heendeniya, A. Sumper, U. Eicker, The multi-energy system co-planning of nearly zero-energy districts – status-quo and future research potential, Appl. Energy 267 (2020) 114953, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261920304657.

[68] W.C. Schoonenberg, I.S. Khayal, A.M. Farid, A Hetero-Functional Graph Theory for Modeling Interdependent Smart City Infrastructure, Springer, Berlin, Heidelberg, 2019, p. 196, [Online]. Available: http://dx.doi.org/10.1007/978-3-319-99301-0.

[69] M. Kivelä, A. Arenas, M. Barthelemy, J.P. Gleeson, Y. Moreno, M.A. Porter, Multilayer networks, J. Complex Netw. 2 (3) (2014) 203–271.

[70] A.M. Farid, D.J. Thompson, W. Schoonenberg, A tensor-based formulation of hetero-functional graph theory, Sci. Rep. 12 (1) (2022) 18805.

[71] 2022. [Online]. Available: https://openmodelingfoundation.github.io/about/.

[72] SE Handbook Working Group, Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities, International Council on Systems Engineering (INCOSE), 2015, pp. 1–386.

[73] D. Hoyle, ISO 9000 Pocket Guide, Butterworth-Heinemann, Oxford; Boston, 1998, p. 496, [Online]. Available: http://www.loc.gov/catdir/toc/els033/99163006.html.

[74] A.M. Farid, D.C. McFarlane, A development of degrees of freedom for manufacturing systems, in: IMS'2006: 5th International Symposium on Intelligent Manufacturing Systems: Agents and Virtual Worlds, Sakarya, Turkey, 2006, pp. 1–6, [Online]. Available: http://engineering.dartmouth.edu/liines/resources/Conferences/IEM-C02.pdf.

[75] A.M. Farid, Reconfigurability Measurement in Automated Manufacturing Systems (Ph.D. thesis), University of Cambridge Engineering Department Institute for Manufacturing, 2007, pp. 1–302, [Online]. Available: http://engineering.dartmouth.edu/liines/resources/Theses/IEM-TP00.pdf.

[76] A.M. Farid, D.C. McFarlane, Production degrees of freedom as manufacturing system reconfiguration potential measures, Proc. Inst. Mech. Eng. B 222 (B10) (2008) 1301–1314, [Online]. Available: http://dx.doi.org/10.1243/09544054JEM1056.

[77] A.M. Farid, Product degrees of freedom as manufacturing system reconfiguration potential measures, Int. Trans. Syst. Sci. Appl. 4 (3) (2008) 227–242, [Online]. Available: http://engineering.dartmouth.edu/liines/resources/Journals/IEM-J04.pdf.

[78] A.M. Farid, Static resilience of large flexible engineering systems: Axiomatic design model and measures, IEEE Syst. J. PP (99) (2015) 1–12, [Online]. Available: http://dx.doi.org/10.1109/JSYST.2015.2428284.

[79] A.M. Farid, An engineering systems introduction to axiomatic design, in: A.M. Farid, N.P. Suh (Eds.), Axiomatic Design in Large Systems: Complex Products, Buildings & Manufacturing Systems, Springer, Berlin, Heidelberg, 2016, pp. 1–47, [Online]. Available: http://dx.doi.org/10.1007/978-3-319-32388-6.

[80] D. Thompson, P. Hegde, W.C.H. Schoonenberg, I. Khayal, A.M. Farid, The hetero-functional graph theory toolbox, 2020, arXiv:2005.10006.

[81] A.M. Farid, D. Thompson, P. Hegde, W. Schoonenberg, A tensor-based formulation of hetero-functional graph theory, 2021, arXiv preprint arXiv:2101.07220.

[82] D. Thompson, A.M. Farid, Reconciling formal, multi-layer, and hetero-functional graphs with the hetero-functional incidence tensor, in: IEEE Systems of Systems Engineering Conference, Rochester, NY, 2022, pp. 1–6.

[83] L.A.N. Amaral, A. Scala, M. Barthelemy, H.E. Stanley, Classes of small-world networks, Proc. Natl. Acad. Sci. 97 (21) (2000) 11149–11152.

[84] R. Albert, H. Jeong, A.-L. Barabási, Error and attack tolerance of complex networks, Nature 406 (6794) (2000) 378–382.

[85] M. Barthélemy, Spatial networks, Phys. Rep. 499 (1) (2011) 1–101.

[86] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (5439) (1999) 509–512.

[87] D.J. Thompson, The American multi-modal energy system: Model development with structural and behavioral analysis using hetero-functional graph theory, 2023.