Engineering 65 (2026) 100–111

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
Energy Transition's Technical Pathways, Risks, and Equity—Article

Algorithm for Transportation Pathways and Patterns Through Pipeline Network: A Case Study in California for the Power Generation Sector

Zemin Eitan Liu [a,*], Diego Moya [b], Zhenlin Chen [c], Wennan Long [a], Liang Jing [b], Bo Ren [d], Haoyu Tang [a], Muhammad Y. Jabbar [b], Farah Ramadan [e], James Littlefield [e], Mohammad S. Masnadi [a,*]

[a] Chemical and Petroleum Engineering Department, University of Pittsburgh, Pittsburgh, PA 15213, USA
[b] Energy Sustainability Analysis Team, Technology Strategy & Planning Department, Dhahran 34465, Saudi Arabia
[c] Energy Science Engineering, Stanford University, Stanford, CA 94305, USA
[d] Global Technology Analysis Team, Aramco Research Center-Houston, Aramco Americas, Houston, TX 77084, USA
[e] Global Technology Analysis Team, Aramco Research Center-Novi, Aramco Americas, Novi, MI 48377, USA

ARTICLE INFO

Article history:
Received 28 May 2025
Revised 3 November 2025
Accepted 10 November 2025
Available online 14 November 2025

Keywords:
Transmission pipeline network
Natural gas
Energy transmission
Pathway planning and optimization
Geospatial supply-chain traceability
Power generation

ABSTRACT

Decarbonization and energy transition in modern energy systems require integrated tools that can analyze complex supply pathways, optimize infrastructure choices, and evaluate policy impacts. This study presents the algorithm for transportation pathways and patterns through pipeline network (ATP3), an open-source framework combining high-resolution supply-chain traceability, cost- and emissions-aware network optimization, and scenario simulation. ATP3 reconstructs real-world flows by matching supply and demand nodes across an entire pipeline network, integrating the entropy weight method (EWM)-based allocation and a minimum-cost flow formulation with scenario-driven computational simplification. In a California power generation case study, ATP3 accurately identified 154 processing-to-plant supply routes and 134 upstream field linkages. In addition, when using the EWM-based 40% baseline allocation, the model determined that 90.53% of natural gas used for electricity was imported—primarily from Arizona (48.66%), Oregon (27.43%), and Nevada (14.44%)—with Arizona supplying the largest single external volume (1034.51 MMscf·d^{-1}, here MMscf is million standard cubic feet, and 1 MMscf ≈ 28 317 m^3). In-state production accounted for only 201.37 MMscf·d^{-1} (9.47%). Meanwhile, EWM allocation reduced transportation costs by 2.76% compared to uniform allocation by favoring geographically proximate sources. These results demonstrate the ability of ATP3 to bridge granular infrastructure mapping with system-level planning, offering a robust and versatile platform for life-cycle assessment, infrastructure planning, and policy evaluation across power, transportation, and industrial sectors. The continuously updated resource is available via the GitHub repository.

© 2025 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Pipeline transportation plays a pivotal role in the natural gas supply chain, with implications for climate change mitigation, energy security, and infrastructure planning. Accurately estimating and optimizing natural gas transmission pathways is essential for quantifying life-cycle greenhouse gas (GHG) emissions [1] and for enhancing operational efficiency [2] and identifying opportunities for a transition toward the adoption of renewable energy systems [3,4]. Previous studies have shown that the GHG intensity of natural gas can vary widely depending on the production site and mode of transportation [5]. Supply disruptions and shifting trade patterns further underscore the need for high-fidelity gas transportation models to inform policy and investment [6,7]. However, most natural gas transportation models suffer from three major limitations that constrain their realism and usefulness: low spatial precision with unrealistic routing assumptions, coarse spatial granularity, and a limited geographic scope. Below, these gaps

* Corresponding authors.
E-mail addresses: eliuzm@163.com (Z.E. Liu), m.masnadi@pitt.edu (M.S. Masnadi).

https://doi.org/10.1016/j.eng.2025.11.005
2095-8099/© 2025 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

are reviewed and motivate the development of a new approach to modeling natural gas pathways.

1.1. Spatial precision and routing assumptions

Several models have simplified pipeline transportation by using straight-line distances or centroid-to-centroid approximations, rather than the actual lengths and routes of pipelines. For example, life-cycle assessments (LCAs) often assume a fixed pipeline distance or use great-circle distances between regional centroids. In a United States analysis, Littlefield et al. [1] used straight-line distances between state-level supply and demand centroids and then applied a 9% distance increase to account for pipeline meandering. In addition, models by Welder et al. [8], Welder et al. [9], Samsatli et al. [10], and Moreno-Benito et al. [11] aggregate spatial resolution into a few large regions and utilize distance matrices for transport between regions. Such approaches have improved completely aggregated national models but still introduce error—actual pipelines seldom follow direct paths, and route length may vary substantially with terrain, permitting, and right-of-way constraints. Data on real pipeline distances have historically been scarce, leading to these crude assumptions. As a result, energy use and emissions from compression stations are often underestimated, and delivery timing and pressure requirements are often misrepresented. In summary, low spatial precision in routing obscures the true impact and feasibility of gas transportation.

Recent studies have highlighted the importance of using real pipeline infrastructure data for modeling. Roman-White et al. [12] found that incorporating detailed pipeline information could significantly alter GHG estimates for liquefied natural gas (LNG) supply chains. The refined routing led to a 22%–53% lower GHG intensity for delivered LNG in their reference case compared to earlier US Department of Energy estimates. These findings underscore the fact that the unrealistic routing assumptions of prior models can result in misleading results and that using actual pipeline routes yields a more accurate and nuanced picture. In practice, a high-fidelity routing approach may reveal multiple distinct pathways for gas flows where simpler models assume only one, thereby capturing the operational flexibility of a network that was previously missed.

1.2. Granularity of infrastructure representation

A related limitation in existing gas transportation models is their coarse spatial granularity. Many frameworks have aggregated infrastructure to large zones or corridors, such as modeling only state-to-state flows [13] or basin-to-market links [12], rather than representing point-to-point connections at the individual-pipeline level. For example, the Natural Gas Market Module, as the component of the National Energy Modeling System (NEMS) representing the North American natural gas transmission system, groups a pipeline network into a limited set of corridor links between supply and demand regions, sacrificing detail for computational tractability [13]. Because of this aggregation, important network features (e.g., specific bottlenecks or alternate routes) are lost, and even flow direction reversals or spot capacity issues cannot be adequately analyzed in those models. High granularity is crucial for capturing the topology and physics of pipeline networks. As Müller-Kirchenbauer et al. [14] noted, detailed spatial disaggregation is needed when analyzing energy infrastructures, as low-resolution models may be insufficient for representing the complexity and interconnectedness of real networks. In practice, coarse models often allocate flows based on simple rules (proximity or regional production share) without an explicit network flow balance [1,12]. This can lead to counterintuitive or infeasible results—for example, delivering more gas into a region than pipeline capacity allows, or neglecting the role of transit states in between.

Increasing the spatial granularity of models enables flow balancing and capacity-constrained routing at the infrastructure level, ensuring that supply and demand are matched through actual pipelines without violating constraints. Fine-grained network models can identify pipelines being utilized, how pipeline pressure limits and compressor stations affect throughput, and where congestion might occur. They also preserve the distinction between multiple parallel paths and entry points. Studies have begun moving in this direction: Lochran [6] introduced the Gas Network Optimisation Model for Europe (GNOME), a highly granular optimization that explicitly represents the pipelines and storage in each country. Their work was motivated by the observation that most European gas models were too coarse to assess infrastructure investments and supply dynamics. Notably, GNOME represents 40 countries (European Union (EU)-27 plus neighbors) as nodes with 242 pipeline and LNG shipping connections, solving flows on a monthly basis. Likewise, Di Lullo et al. [15] developed a bottom-up GHG assessment for gas pipelines, acknowledging that a detailed, segment-level approach was necessary to accurately quantify emissions and energy use in transmission systems. These examples reveal a clear trend toward higher-resolution modeling. Still, even the most advanced regional models can be limited by data availability and computational complexity. GNOME aggregates all intra-country infrastructure to one or two nodes per country, and NEMS, despite modeling hundreds of pipeline corridors, uses a simplified representation of pipeline networks connecting nodes that represent aggregated regions. Achieving fine granularity without losing computational efficiency or transparency remains a challenge.

1.3. Geographic scope and system boundaries

Beyond spatial detail, geographic scope is a critical factor for realistic gas pathway modeling. In a number of existing models, analyses are limited to intra-regional transport, ignoring cross-border or interstate gas deliveries that, in reality, are commonplace. This limitation can introduce significant biases. For example, California consumes large amounts of natural gas but produces very little; it obtains 90% of its supply from other regions via interstate pipelines [16]. If a model analyzed California's gas system without accounting for those imports, it would misidentify supply routes and likely underestimate transport distance and the associated emissions. Burns and Grubert [5] addressed this type of problem by mapping production emissions to consumption locations. They found that the methane intensity of delivered gas varied considerably by state (0.9%–3.6%) precisely because some states rely on long-distance imports from high-leakage production basins while others use more local supply. This nuance would be lost in single-region analyses. Similarly, many European gas security models historically examined only intra-EU flows or a subset of countries, neglecting the role of external suppliers [6]. This could lead to an underestimation of the dependency on external infrastructure, as well as an incomplete view of supply resilience.

Expanding the geographic scope to include full supply routes would yield more reliable and unbiased results. LCAs should include all relevant transport steps (from the point of production to the point of use, including all midstream segments) to accurately assign emissions and energy use. Littlefield et al. [1] emphasized that a connected supply-chain perspective is necessary for differentiating high- and low-emission gas sources, something not possible with isolated regional LCAs or national averages. In terms of energy security and planning, a wide scope allows models to capture interregional flow shifts (e.g., how a supply cut in one region is compensated by imports from another) and to identify infrastructure that lies outside a focal region but is critical for its supply.

Another shortcoming of traditional gas transport models is their reliance on predefined supply-to-demand pairings. Often, models simply assign each demand center a primary supply source (or a fixed fraction from each source) based on historical patterns or a cost ranking, ignoring the actual network routing or capacity constraints. This approach can significantly skew results, especially in dynamic markets or under shifting policy regimes where supply sources and demand centers are dynamic. Similarly, relying on historical pairing patterns alone does not account for the potential operational flexibility of pipeline networks or economic rerouting that could occur under new conditions. In reality, gas networks can respond to price signals and capacity availability by rerouting volumes through alternate pipelines. Failing to model this flexibility means missing important behaviors; for example, during a supply disruption, gas might travel a longer route through a neighboring country, or previously underutilized pipelines may become critical pathways. In summary, the limited geographic scope and hardwired flow assumptions of many models restrict their ability to analyze system-wide gas dynamics and stress-test networks under transformative scenarios.

1.4. Contributions of the present work

To address the lack of spatial precision, infrastructure granularity, and transparent scope in existing natural gas transportation models, this study proposes the algorithm for transportation pathways and patterns through pipeline networks (ATP3), a modeling framework that integrates infrastructure-resolved geospatial data with an optimization-based network formulation. Detailed comparisons of ATP3 and the most widely used energy transport models are shown in Table 1 [6,13,17–21]. ATP3 was designed to trace natural gas flows from supply points to demand points through a pipeline network, under a clearly defined steady-state or snapshot scenario. This steady-state focus aligns with the overarching objective of building a robust baseline for subsequent analyses of midstream-related carbon footprints, emissions, and energy consumption.

Furthermore, ATP3 was developed to overcome the persistent challenges in achieving infrastructure-level traceability, optimizing flow allocation under realistic network constraints, and quantifying sector-specific supply patterns in large-scale natural gas systems. Unlike prior approaches, it traces point-to-point pathways along actual pipeline infrastructure and allocates flows optimally across an entire network, thereby preserving the physical fidelity of route lengths and enabling accurate estimations of distance-dependent losses and emissions. By integrating publicly available pipeline datasets into a graph-theoretic network representation, ATP3 operates at the resolution of individual segments and nodes, bridging the granularity gap in existing models. Its formulation incorporates a complete system boundary, capturing cross-regional deliveries, including flows across state and national borders, to connect each supply source with each demand point. The optimization framework specifically matches supply with demand based on current network topology and facility conditions,

Table 1
Comparison of selected energy transport models in terms of spatial precision, infrastructure granularity, and scope.

Model framework	Spatial routing precision	Infrastructure granularity	Geographic/system scope	Temporal/scenario capability
MARKAL/TIMES [17]	Great-circle or average-distance routing; no explicit pipeline paths (✗)	Detailed techno-economic representation for plants, but aggregated transport (✓)	Global or multi-region, but trade flows are aggregated; no detailed networks (✓ scope, ✗ network)	Long-term scenarios (multi-decade) via optimization (✓ temporal)
GCAM/MESSAGE [18]	Inter-region trade only; highly simplified distances (✗)	No specific pipelines; aggregated infrastructure (✗)	Global with large multi-country regions (✓ global scope)	Long-term scenario projections (✓ temporal)
PRIMES [19]	Aggregated corridors between European zones; no real geospatial routes (✗)	Detailed tech options for supply/demand, but gas transport aggregated (✗)	European multi-country zones; cross-border flows modeled abstractly (✓ regional scope)	Multi-year policy scenarios for EU (✓ temporal)
NEMS [13]	State-level nodes with pipeline corridor distances; some meandering accounted (✓)	Represents ~121 nodes and 530 pipeline links in North America; major pipelines and storage included, but aggregated within states (~)	National (US + limited Canada/Mexico) with import/export points; not global (~)	Annual projections to ~2050; policy scenario analysis in Annual Energy Outlook (✓ temporal)
GNOME [6]	Nodal network representation; uses real gas line data when available (✓)	Explicit country-to-country network with pipelines, LNG ports, storage (✓); each country often a single node (limits intra-country detail) (~)	European system (EU-27 + neighbors) with external suppliers modeled; not worldwide (~)	Dynamic multi-period (monthly, 2025–2040) optimization; can test disruption scenarios (✓ temporal)
METIS [20]	Realistic routing for power & gas networks in EU (✓)	Mixed granularity: some regional aggregation remains for gas (~)	EU-wide with import nodes for external sources (✓ regional scope)	Hourly power/gas dispatch simulations; used for EU scenario studies (✓ temporal)
HDSAM [21]	User-defined point-to-point routing; pipeline length input by user (✓ single-path)	Detailed components (compressors, liquefiers, etc.) for hydrogen delivery costs (✓ component detail); evaluates one delivery pathway at a time (✗ network flow)	Project- or city-scale hydrogen supply scenarios; not an integrated energy system model (✗ broad scope)	Static analysis of given scenario; no multi-period evolution (✗ temporal)
ATP3 (this work)	GIS-based routing along actual pipeline corridors; preserves true path lengths and topology (✓)	Node-level asset modeling: individual pipelines, compressors, storage, power plants connected in network (✓)	Flexible scope from regional to global; traces supply routes across borders with source-to-sink fidelity (✓)	Current version is steady-state; can be iterated for scenarios, but transient or multi-period dynamics not explicit (~)

Legend: ✓: explicitly supported; ~: partially supported; ✗: largely absent/simplified.
Methodological note: Spatial routing precision refers to whether models use real mapped routes versus abstract distances. Infrastructure granularity indicates the level of detail in representing pipelines and nodes (individual assets vs aggregated links). Geographic/system scope denotes the breadth of the modeled network (sub-national, national, global) and whether external supply/demand are represented. Temporal/scenario capability reflects the ability to simulate changes over time or test policy scenarios. Assessments (✓/~/✗) focus on gas transmission and are based on published documentation; some models excel outside gas pipelines (e.g., HDSAM's detailed hydrogen components).
MARKAL: market allocation; TIMES: the integrated MARKAL–energy flow optimization model (EFOM) system; GCAM/MESSAGE: global change assessment model/model for energy supply strategy alternatives and their general environmental impact; METIS: model for energy supply systems and alternative technologies for long-term planning; HDSAM: hydrogen delivery scenario analysis model; PRIMES: price-induced market equilibrium system; GIS: geographic information system.

accommodating capacity and operational constraints while remaining extensible to cost, emission, or energy-use objectives. Together, these capabilities allow ATP3 to provide physically grounded, high-resolution simulations of natural gas transportation patterns that address the critical need for accurate, infrastructure-resolved flow modeling in energy systems research.

ATP3 comprises a graph-theoretic network representation of pipelines, nodes, and physical connectivity constructed from real geospatial data; node-level supply and demand definitions derived from publicly available infrastructure and sectoral datasets; and a flow optimization core formulated as a minimum-cost flow (MCF) problem that can be solved directly using standard solvers. In addition, the model incorporates a proportion allocation scheme for supply–demand matching, which applies an entropy weight method (EWM) to objectively determine facility-specific allocation ratios based on multiple attributes when detailed sectoral data are unavailable. The main contributions of this work are as follows:

(1) Novel integration of EWM into supply–demand balancing—by introducing EWM for allocation ratio estimation, ATP3 provides a systematic, data-driven approach to handling incomplete sectoral information and to differentiating supply contributions among facilities according to their attributes.

(2) Flexible MCF-based formulation with computational simplification—ATP3 frames the pipeline transportation problem as an MCF model and, where justified by scenario constraints, simplifies it to a shortest-path and transportation problem approach. This yields results mathematically equivalent to the full MCF under the given assumptions but with a substantially reduced computational cost.

(3) Fine-grained case study for California's power generation sector—applying ATP3 to California enables a high-resolution analysis of the state's natural gas sourcing patterns for electricity generation, tracing individual pathways from production to consumption and identifying cross-regional delivery patterns.

By combining infrastructure-level geospatial data, an adaptable optimization core, and a proportion allocation framework, ATP3 provides both methodological advances and practical applicability for sector-specific, cross-regional, point-to-point continuous natural gas flow analysis—capabilities that are not simultaneously present in existing coarse-scale models.

The remainder of this article is organized as follows. Section 2 introduces the core methodology and workflow of ATP3, with detailed functions and scope provided in Appendix A. Section 3 presents a California-based case study to demonstrate the capabilities of ATP3 and validate its performance. Section 4 discusses the key takeaways, implications, and limitations of ATP3. Section 5 concludes the article.

2. Methodology and workflow formulation

Fig. 1 illustrates the workflow for the case study implementation using ATP3. The workflow begins with the preprocessing of input data. For real-world scenario simulations, data should be collected following the format specified in Appendix A. Alternatively, the framework allows for hypothetical data inputs, either to represent synthetic case studies or to address gaps in data availability, ensuring that the input set remains complete and computationally tractable. Following this, a network graph is constructed from the processed data. The next step is to assign weights to each edge based on different objectives and balance the supply–demand relationship. If applicable, an optimization procedure is then applied to determine allocation strategy. Finally, results are visualized and interpreted.

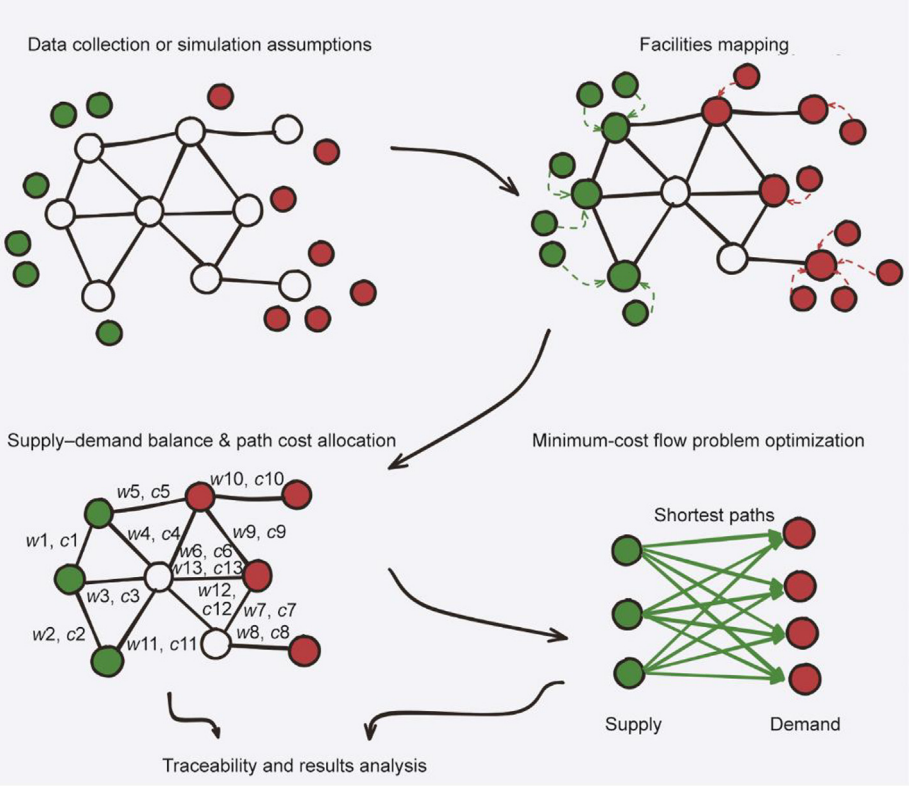


Fig. 1. Workflow of ATP3 for case study. Green nodes denote supply facilities; red nodes denote demand facilities. Black lines are the pipeline segments with the junctions. Green arrow lines are the shortest weighted transmission pathways. *w*1–*w*13: transportation features of edge 1–13; *c*1–*c*13: transportation capacities of edge 1–13.

2.1. Graph construction by data

According to the pipeline data, the network is represented by a weighted graph $\mathscr{G} = (\mathscr{N}, \mathscr{E}, \mathscr{W})$, where $\mathscr{N}$ is the set of nodes and $\mathscr{E}$ is the set of pipeline edges. $\mathscr{W}$ represents the weight of each edge connecting two nodes, including the transportation cost per unit volume of natural gas (detailed in Section 2.3) and the maximum transportation capacity of the pipeline segment.

2.2. Facilities mapping

Real-world facilities, specifically production fields ($\mathscr{F}$), processing plants ($\mathscr{P}$), and terminals ($\mathscr{T}$), are mapped onto the nodes ($\mathscr{N}$) of $\mathscr{G}$ so that they become part of the node set ($\mathscr{F} \subseteq \mathscr{N}$, $\mathscr{P} \subseteq \mathscr{N}$, $\mathscr{T} \subseteq \mathscr{N}$). Each field (supply source), plant (intermediate processing facility), and terminal (delivery or demand point) is associated with a specific node in the graph. If a facility is physically located at an existing pipeline node, it is directly mapped to that node. Otherwise, the facility is linked to the nearest pipeline node by adding a connector edge representing a short feeder pipeline. As a result, any route from a field to a terminal will correspond to a path through graph $\mathscr{G}$ (potentially passing through a plant node if processing is required).

2.3. Supply–demand balance and path cost allocation

In most cases, supply and demand are imbalanced owing to insufficient data coverage or limitations in data quality. To achieve a balanced allocation of point-level supply–demand transportation, this article proposes a multi-step procedure based on EWM. For clarity, the method is illustrated using a case in which the total demand is lower than the available supply; the converse scenario can be addressed analogously. In this example, the supply proportion for each supply point must be determined. This method first calculates a set of key operational indicators for each point and applies normalization to account for differences in scale and directionality. Subsequently, the normalized indicators are transformed into proportional form across all supply points, yielding a relative performance matrix. The information entropy for each indicator is then computed as:

$$E_{\mathrm{EWM},j} = -\frac{1}{\ln n}\sum_{i=1}^{n} p_{\mathrm{EWM},ij}\ln(p_{\mathrm{EWM},ij} + \varepsilon) \tag{1}$$

where $p_{\mathrm{EWM},ij}$ is the proportion of indicator j for supply point i, n is the total number of supply points, and ε is a small constant to avoid singularity. Indicator divergence values were calculated as $d_{\mathrm{EWM},j} = 1 - E_{\mathrm{EWM},j}$, and the entropy-based weights were obtained by:

$$w_{\mathrm{EWM},j} = \frac{d_{\mathrm{EWM},j}}{\sum_{m=1}^{M} d_{\mathrm{EWM},m}} \tag{2}$$

where m is the indicator and M is the total number of indicators. The composite EWM score for each supply point is then computed as:

$$s_{\mathrm{EWM},i} = \sum_{j=1}^{M} w_{\mathrm{EWM},j}\, p_{\mathrm{EWM},ij} \tag{3}$$

These scores are scaled such that the weighted total allocated supply matches a predefined proportion of the aggregate available flow, while ensuring that no point exceeds its supply capacity. The resulting allocation ratio for each supply point (ρ_i) is calculated by:

$$\rho_i = \frac{\min(s_{\mathrm{EWM},i} \cdot B \cdot \sum_{i=1}^{n} Q_i)}{Q_i} \tag{4}$$

where $\sum_{i=1}^{n} Q_i$ is the total supply flow, B denotes the predefined baseline proportion allocated to demand. ρ_i is subsequently used as the transportation weight for its outgoing flows in the network model, ensuring a consistent and capacity-constrained supply–demand balance across the system.

After reconciling supply and demand at the point level, the network edges are assigned weights that capture both transportation features and physical constraints. Specifically, the weight of each pipeline segment is decomposed into two components: ① a transportation feature (w), representing the transport cost, emissions, or energy consumption per unit volume of natural gas transported, and ② a transportation capacity (c), denoting the maximum allowable throughput for that segment.

In many existing studies, cost, emissions, and energy usage were well-approximated as linear functions of both the transported volume and the transportation distance [22–24], allowing these features to be embedded within a distance-based weight metric. Accordingly, in this study, w was simplified to be the pipeline length, which could subsequently be scaled by target-specific coefficients to reflect alternative optimization objectives.

To determine the pipeline length (d_{ij}) of each pipeline segment (i,j), the Haversine formula [25], which is widely used for geodesic distance calculations owing to its numerical stability and accuracy, even under extreme conditions such as when points are extremely close or cross the International Date Line, was adopted as:

$$\begin{cases} \alpha = \sin^2\left(\frac{\Delta \mathrm{lat}}{2}\right) + \cos\left(\mathrm{lat}_i\right)\cos\left(\mathrm{lat}_j\right) + \sin^2\left(\frac{\Delta \mathrm{lon}}{2}\right) \\ \beta = 2 \cdot \arctan(\sqrt{\alpha}, \sqrt{1-\alpha}) \\ d_{ij} = R_{\mathrm{Earth}} \cdot \beta \end{cases} \tag{5}$$

where α is an intermediate variable in the haversine formula, β is the central angle between two geographic points on the Earth's sphere, lat_i is the latitude of node i on the Earth's surface, Δlat and Δlon denote the differences in latitude and longitude between nodes i and j, respectively, and R_{Earth} is the Earth's radius (set to 3960 m in this study).

2.4. MCF problem optimization

Following the supply–demand balance stage and path cost allocation, the optimization of transmission traceability and patterns across the entire network can be naturally formulated as an MCF problem. The mathematical formulation and solution procedures are presented in detail in Appendix A. In the specific context of this study, the volumetric natural gas flow is significantly lower than the transportation pipeline capacity [26], which allows capacity-limitation effects to be neglected without compromising the validity of the results. Under this assumption, this article further postulates that gas between any two nodes is transported along the shortest path in the network graph ($\mathscr{G}$), as determined by the pre-computed edge weights ($\mathscr{W}$). This assumption is reasonable in the present context because, in the absence of binding capacity constraints, shortest-path routing minimizes distance-dependent metrics such as transportation cost, energy usage, or emissions, consistent with prior findings in pipeline optimization studies [27].

These considerations simplify the MCF problem into two sequential subproblems: ① determining the shortest paths between all supply–demand pairs, and ② solving a transportation problem over the resulting graph. This reduction considerably lowers computational complexity compared to solving the full MCF on the original network, as it eliminates intermediate node flow variables and leverages specialized transportation problem solvers that scale efficiently to large network systems [28]. Moreover, it facilitates scenario analysis by allowing cost coefficients to be modified directly in the reduced cost matrix without re-running full network flow optimizations.

(1) Shortest-path calculation and path weights. With a weighted graph representation in place, shortest-path algorithms

were employed to evaluate distances and viable routes between facilities. In particular, Dijkstra's algorithm [29] was adopted as an example to find the shortest path between nodes in the weighted pipeline graph. The process is shown as follows:

- Initialize all nodes with infinite distance, except the source node, which is set to zero.
- Repeatedly select the unvisited node with the smallest distance, consider all its unvisited neighbors, and calculate their tentative distances.
- Mark the current node as visited and continue with the node that has the smallest tentative distance.
- Continue this process until all nodes have been visited.

Dijkstra's algorithm systematically explores the graph to identify the minimum-weighted route between a given source and sink, considering the cumulative weights of edges along each path. Using this algorithm, ATP3 computes the shortest path length between every single pair. These shortest-path distances serve as the effective transportation costs for moving products between those locations via the pipeline network. By pre-calculating shortest paths, the model identifies the most efficient route for any given supply–demand pair in the network. These route calculations form the basis for both the traceability module and the definition of distance-based costs in the optimization module.

(2) Transportation problem optimization model. In transportation problem formulation, the decision variables (x) are defined as the amount of product transported from each source to each sink through the network. In practice, because flows may pass through intermediate plants, the model accounts for multi-segment paths. For example, a field may send a product to a specific plant, which then sends it onward to a terminal—constituting a field–plant–terminal route. The objective function J minimizes the total transportation cost, which is typically the distance-weighted volume of the product moved, as shown in Eq. (6).

$$\min: J = \sum_{i,j} w_{ij} x_{ij} \tag{6}$$

where w_{ij} represents the cost of transporting gas from source i to sink j, and x_{ij} is the capacity of gas transported. This objective reflects a desire to minimize the total pipeline usage distance or, equivalently, to minimize the transportation time or cost.

Subject to this objective, the model imposes several physical constraints. Supply constraints ensure that each source does not ship more volume than its production or processing capacity. Demand constraints ensure that each terminal's required volume is met. The constraints can be shown as follows:

$$\text{s.t.} \begin{cases} \sum_j x_{ij} \le S_i, \forall i \\ \sum_i x_{ij} \le D_j, \forall j \end{cases} \tag{7}$$

where S_i represents the supply of node i; D_j denotes the demand of node j.

In summary, the optimization algorithm decides on how much each source should send to each sink in order to satisfy all demands without exceeding capacities, while minimizing the total distance-weighted transportation cost. As demonstrated above, when analyzing a steady-state snapshot of the system, the problem can be optimized by linear programming [30] to obtain a globally optimal solution.

3. Case study and results: California natural gas pathways to power plants

3.1. Case study context and policy relevance

California is an ideal region for demonstrating the capabilities of ATP3, given its large gas-fired power fleet, heavy reliance on imported natural gas, and ambitious GHG reduction goals. The state currently imports roughly 90% of its natural gas from out-of-state basins via interstate pipelines [31], underscoring a complex supply network feeding local power plants. Meanwhile, California has committed to a 100% carbon-free electricity supply by 2045 [32]. This policy context makes supply-chain transparency and emissions accounting critical: Policymakers need to know where power-plant gas is coming from and what the upstream carbon footprint is.

This article leverages publicly available data on power plants, gas processing facilities, and gas fields (including their coordinates and capacities) to trace energy flows. Power generation data and gas processing data were sourced primarily from the Homeland Infrastructure Foundation Level Database [33] in 2022 and the US Energy Information Administration (EIA) [34] in 2017, respectively. Upstream production parameters were calculated using the oil production GHG emissions estimator (OPGEE) model, combined with detailed operational data from multiple authoritative sources [35–37]. Pipeline network data were obtained primarily from the National Renewable Energy Laboratory [38].

According to the data adopted, in the year 2022, there were 285 power plants with a consumption capacity of 2125.95 million standard cubic feet per day (MMscf·d^{-1}; 1 MMscf ≈ 28 317 m^3), 18 processing plants with a processing capacity of 599.95 MMscf·d^{-1}, and 87 fields with a production capacity of 393.68 MMscf·d^{-1} in California. According to data from the EIA, California's total natural gas consumption in 2022 reached 5642.08 MMscf·d^{-1}, with approximately 37.68% attributed to the power generation sector. Notably, the net imported capacity reached 5324.17 MMscf·d^{-1}, accounting for approximately 94.37% of the total consumption, underscoring the state's structural dependence on external supply sources.

The facilities used for the California case study are illustrated in Fig. 2 [2]. ATP3 was then applied to map multi-stage fuel pathways, from gas fields and processing plants across the United States to California power generators, providing insights for researchers, regulators, and industry stakeholders alike.

3.2. Proportion determination

In the field-to-processing-plant stage, national-level supply and demand are balanced, allowing the ATP3 optimization module to directly determine transportation pathways and flow patterns throughout the United States network [39]. By contrast, in the processing-plant-to-end-user stage, the total demand is smaller than the total supply due to the absence of explicit pairing data between processing plants and downstream power plants, facility-level information on the proportion of natural gas directed to power generation, and corresponding data for other end users. Under these conditions, the optimization and simulation modules of ATP3 are leveraged to infer plausible allocation patterns.

Specifically, a range of scenarios is simulated under varying assumptions to identify the most reasonable baseline gas proportion from processing plants to power plants, corresponding to B in Eq. (4). In this analysis, the effects of two alternative schemes on the resulting allocation patterns are compared: ① dynamically adjusting the supply proportion of each processing plant using EWM, and ② assuming a uniform distribution ratio across all processing plants. In the EWM framework, two indicators are selected: the total available flow (treated as a positive indicator, as a larger supply capacity is preferable) and the average shortest-path transportation cost from each supply node to its connected demand nodes (treated as a negative indicator, as shorter transport distances are more desirable).

Fig. 3 shows the comparison results of different proportions after optimization. Whether using the EWM-based or the uniform

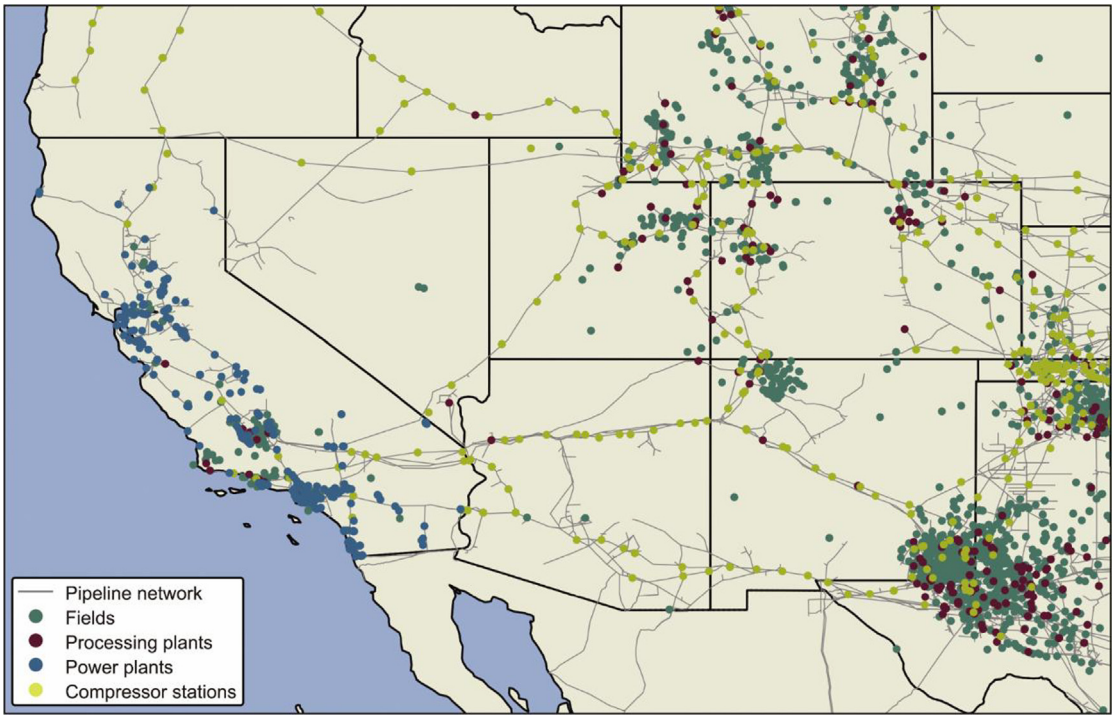


Fig. 2. Potential facilities for California gas supply network map. Given California's heavy reliance on interstate natural gas imports, the map extent was expanded to encompass key upstream supply regions, including the Rocky Mountains, Permian Basin, and San Juan Basin. In addition to fields, processing plants, and power plants, the locations of compressor stations are also highlighted, providing a foundational layer for subsequent analyses of emissions and pipeline pressure behavior [2].

allocation approach, the set of source processing plants and the total number of pathways remained unchanged. This invariance arose because, relative to the uniform method, the EWM approach primarily alters the allocation ratios of processing plants with low total throughput (Fig. 4). Such changes exert minimal influence on the overall network pairing structure or the number of transportation pathways. Contrastingly, substantial shifts in pathway connectivity were driven predominantly by high-capacity processing plants, whose allocations dominated the network flow configuration. Moreover, the total transmission cost was consistently lower under the EWM approach compared with that under the uniform allocation. This reduction arose because EWM assigns relatively higher supply proportions to processing plants located closer to demand nodes (Fig. 5). Such a distance-sensitive allocation shifts

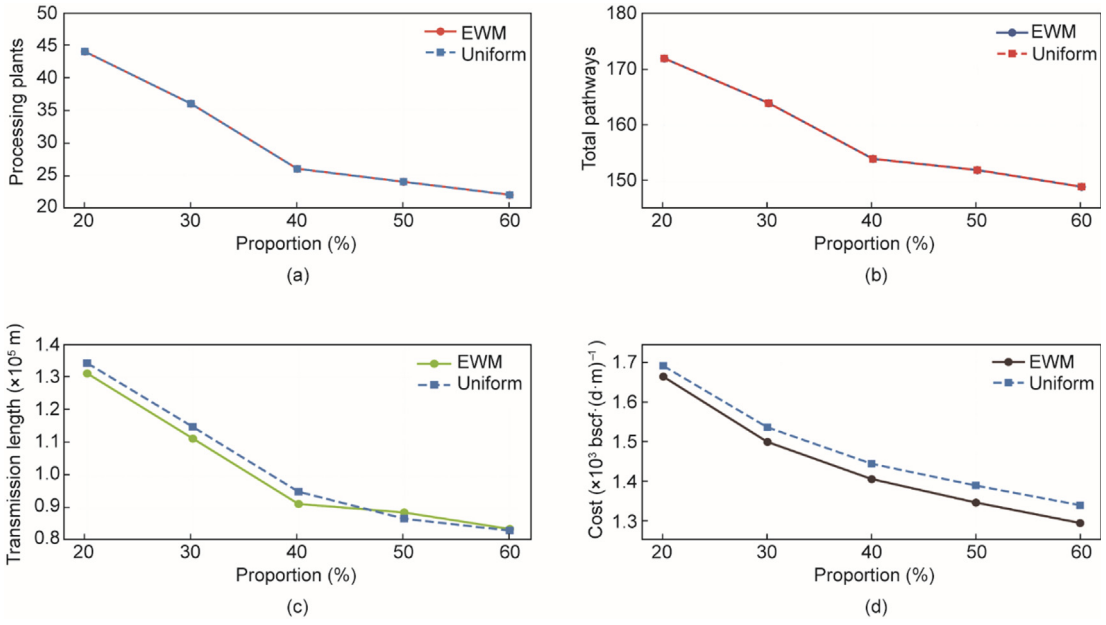


Fig. 3. Comparison of different proportions of natural gas from processing plants supplied to power plants: (a) source processing plant number, (b) total pathway number, (c) total transmission length, (d) total transmission cost. bscf: billion standard cubic feet.

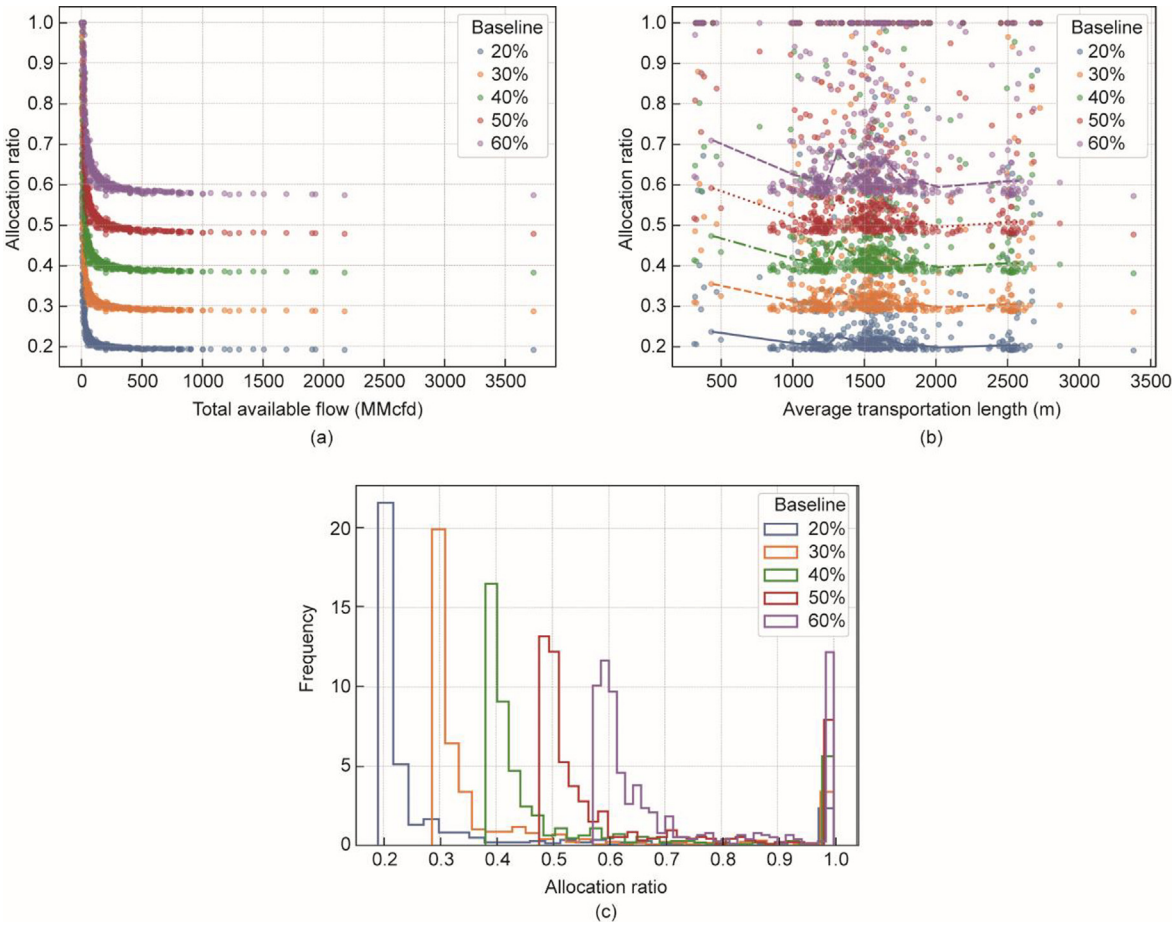


Fig. 4. Baseline-dependent allocation patterns. Comparative analysis of allocation ratios to power plants under five baseline proportions (20%, 30%, 40%, 50%, and 60%). (a) Scatter plots of allocation ratio vs total available flow, showing a clear positive association across baselines. (b) Scatter plots of allocation ratio vs average shortest-path distance, where overall trends are weaker; median-binned trend lines (12 frequency-based bins) are overlaid to highlight subtle correlations. (c) Frequency distributions of allocation ratios, illustrating how different baseline assumptions shift both the spread and the central tendency of allocation outcomes.

the transportation pattern toward shorter routes, thereby reducing the overall transportation cost even when the set of supply–demand pairs and the underlying pathways remain unchanged.

Our findings revealed that the source processing plants, total pathway number, total transmission length, and total transmission cost (defined as the total flow weighted transportation cost, shown in Eq. (6)) exhibited a decreasing trend with increasing baseline proportion for both EWM and uniform allocation. This occurred because directing a larger share of natural gas from each processing plant to the power plants reduced the number of processing plants required as supply sources. Consequently, the overall transportation length and associated transportation cost were reduced owing to the more spatially concentrated supply configuration. However, the decreasing trend became less pronounced once the allocation proportion exceeded 40%. This was because, even under the assumption that all natural gas processed within California is allocated to power generation, the in-state supply would remain insufficient to meet the total demand. As a result, long-distance imports from out-of-state processing plants—such as those in the Rocky Mountains and San Juan Basin—remain necessary. In contrast, when the proportion was below 40%, the shortage was more significant, and supply had to be supplemented from more distant sources, including the Permian Basin and beyond, leading to increased transmission distances and costs.

Taken together, the simulation results indicate that EWM yielded 2.76% lower transportation costs than those of uniform allocation, with a 40% baseline emerging as a robust and representative assumption. This proportion is consistent with the observed share of power sector consumption in empirical datasets and marks a balance point where supply efficiency aligns with geographical constraints. Accordingly, the subsequent analyses and results adopt the EWM-based 40% baseline as the default allocation scenario.

3.3. Transportation patterns: Optimized gas flows to power plants

After establishing the 40% baseline allocation based on EWM and using input data on plant demands, field production, and processing capacity, ATP3's optimization module was applied to determine the optimal flow assignments. This process effectively inferred point-to-point connections that minimize transportation costs while adhering to capacity constraints.

The resulting supply-chain pathways are shown in Fig. 5, where Fig. 5(a) depicts flows from production fields to processing plants and Fig. 5(b) illustrates flows from processing plants to power plants. In this scenario, ATP3 resolved 154 supply routes linking power plants to processing facilities, covering a cumulative transmission distance of 91 475 m, and further identified 134 upstream routes from those facilities to their originating fields, spanning 29 010.3 m. The inferred patterns are consistent with established reports. Many northern California plants are predominantly supplied from the Rocky Mountain region, whereas southern

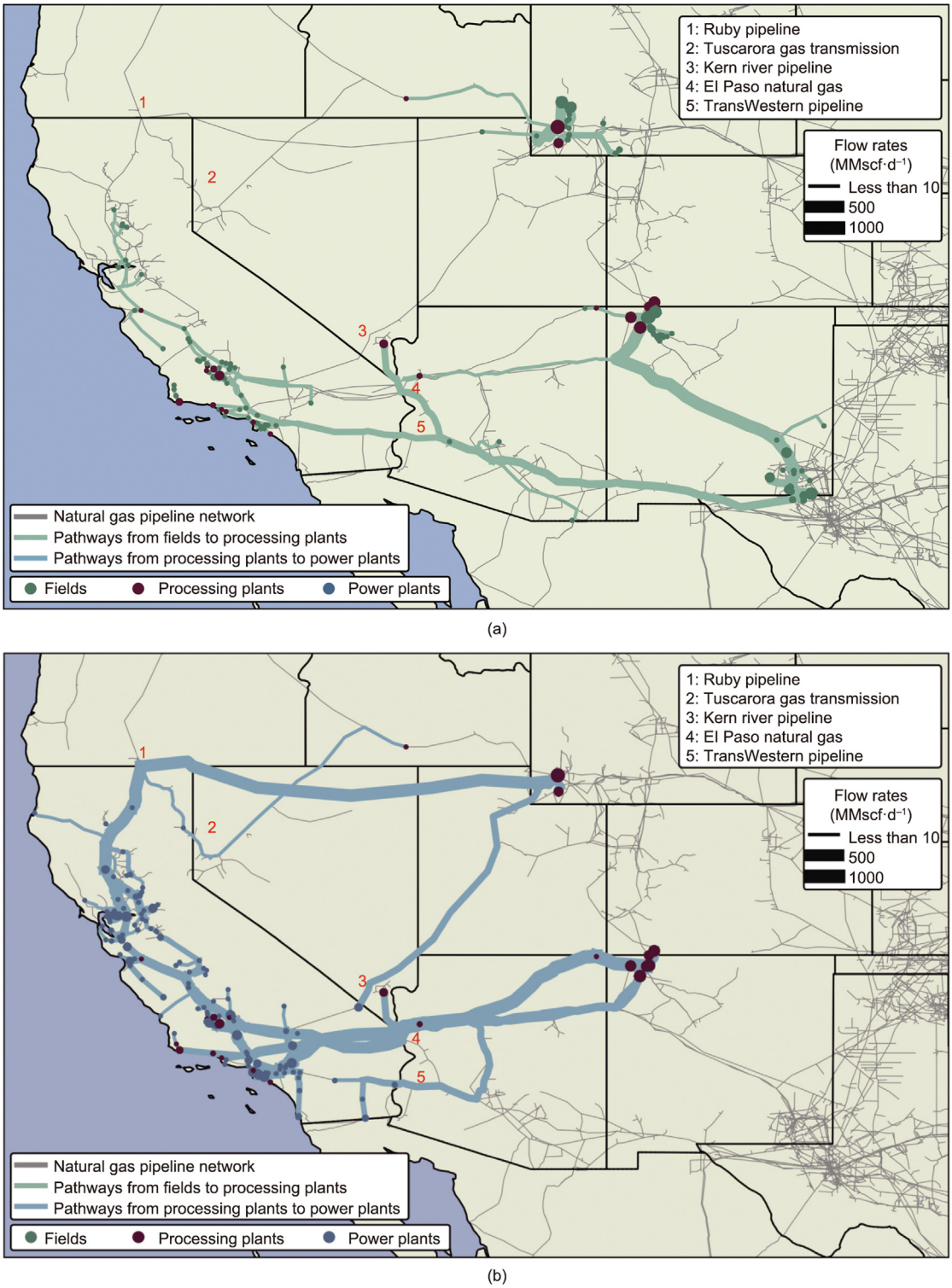


Fig. 5. Natural gas supply pathways to California power plants. The supply network is separated into two key stages: (a) simulated transportation pathways from production fields to processing plants (light green lines), and (b) supply pathways from processing plants to power plants (light blue lines). In both panels, line thickness represents the volumetric flow rates (MMscf·d^{-1}), and node size reflects facility capacity. California's natural gas imports primarily depend on five major interstate transmission pipelines, highlighted by red numeric labels: 1 refers Ruby pipeline, 2 refers Tuscarora gas transmission, 3 refers Kern river pipeline, 4 refers El Paso natural gas, and 5 refers TransWestern pipeline.

California plants receive substantial volumes from the Permian Basin and the broader southwest, an allocation consistent with California's pipeline interconnections to these basins [40].

To validate the model outputs, the supply flow patterns inferred by ATP3, specifically the connections from processing plants to power plants, were compared with the major gas transmission routes known to serve California. The results revealed high consistency with established supply dynamics (Fig. 5(b)). Notably, ATP3 correctly identified the Rocky Mountains and the San Juan Basin as dominant origins for gas supplied to California's power

generation sector after processing. In the model, gas imports from Oregon, Nevada, and Arizona were routed via the Ruby pipeline, Tuscarora pipeline, Kern river transmission, El Paso natural gas, and TransWestern pipeline, mirroring the infrastructure and routing documented in operational reports [16].

Following the qualitative validation of the inferred routes of ATP3 against established gas transmission infrastructure, this part extends the analysis by comparing ATP3's modeled transportation patterns with aggregated statistics from the EIA. Of note, this stage is not intended as a quantitative validation in the strict sense, as the publicly available datasets differ in both spatial granularity and sectoral scope, precluding a point-by-point volumetric match. Instead, the comparison is framed as a consistency check at the sectoral and aggregate levels, with a particular focus on the power generation sector.

Table 2 presents the composition of the natural gas supply to California, disaggregated by source region and sectoral scope. For the power generation sector, Arizona emerged as the dominant external contributor, supplying 1034.51 MMscf·d^{-1}, which accounts for nearly half (48.66%) of the sector's total consumption. By contrast, California's in-state production contributed only 201.37 MMscf·d^{-1} (9.47%), with the remainder sourced from Oregon (583.15 MMscf·d^{-1}, 27.43%) and Nevada (306.92 MMscf·d^{-1}, 14.44%). When expanded to total state consumption across all sectors, Arizona remained the largest supplier at 2165.34 MMscf·d^{-1} (37.87%), followed by Oregon (1861.66 MMscf·d^{-1}, 32.56%) and Nevada (1297.16 MMscf·d^{-1}, 22.69%).

The calculated power-sector share (the ratio of power generation usage to total consumption for each source region) reveals marked differences in allocation patterns. Arizona and California supplied nearly half of their delivered gas to the power sector (47.78% and 51.15%, respectively), while Oregon and Nevada allocated substantially less (31.32% and 23.66%). This disparity suggested that gas imported from Arizona and produced in-state was preferentially directed toward power generation, whereas supplies from Oregon and Nevada were more heavily consumed by other sectors, such as residential, commercial, or industrial users. From a systems perspective, these results highlight two key aspects:

(1) Dependence on interstate imports for power generation—nearly 90% of the gas used for electricity production in California is imported, underscoring the state's vulnerability to upstream supply disruptions, particularly from Arizona.

(2) Source-specific end-user patterns—the strong coupling between Arizona-sourced gas and the power sector may reflect pipeline routing, contractual supply structures, or competitive pricing, whereas Oregon and Nevada imports appear to serve a broader end-use portfolio.

Overall, the combination of model-based flow tracing and independent statistical verification provides a robust foundation for quantifying interstate supply dependencies and their implications for various sectors of energy system planning.

Table 2
Natural gas composition by source regions.

Composition by source region	Scope[a]		
	Power generation sector[b]	Total consumption[c]	Power sector share
California (local)			
Value	201.37[d]	393.68	51.15%
Share	9.47%	6.88%	
Oregon			
Value	583.15	1861.66	31.32%
Share	27.43%	32.56%	
Nevada			
Value	306.92	1297.16	23.66%
Share	14.44%	22.69%	
Arizona			
Value	1034.51[d]	2165.34	47.78%
Share	48.66%	37.87%	
Sum (MMscf·d^{-1})	2125.95	5717.84[e]	37.18%

[a] Value (MMscf·d^{-1}) indicates the absolute quantity of natural gas within the current scope that originates from the given state, whereas share represents the proportion of the total within the current scope contributed by that state.

[b] Based on the ATP3 optimization using the EWM-based 40% baseline allocation, to obtain the value and percentage share of each source region for the power generation sector.

[c] EIA reports total usage (all sectors) at the state level (EIA, 2022); the data are normalized to percentages for structural comparison with the power-only scope.

[d] According to ATP3, the volumes of natural gas imported from Arizona are 929.01 and 206.27 MMscf·d^{-1} in the processing-to-power and field-to-processing stages, respectively. The remaining 306.87 MMscf·d^{-1} in the processing-to-power stage is supplied locally, with Arizona gas accounting for 34.38% of California's field-to-processing stage supply. Assuming this proportion holds for the local supply in the processing-to-power stage, 105.50 MMscf·d^{-1} of the 306.87 MMscf·d^{-1} originates from Arizona. From an LCA perspective, the total volume of gas sourced from Arizona is therefore 1034.51 MMscf·d^{-1}, and the local supply should be 201.37 MMscf·d^{-1}.

[e] Based on EIA's interstate import data combined with the field data used in this article, the resulting total aligns with EIA's reported overall consumption of 5642.08 MMscf·d^{-1}, with a discrepancy of only 1.32%. When accounting for transmission losses and system uncertainties, this deviation is within an acceptable range.

4. Key takeaways, implications, and limitations

Scientific research perspective. The California case study demonstrated that the ATP3 optimization and simulation module can robustly construct real-world infrastructure linkages even when explicit data are incomplete. By leveraging an MCF formulation and point-to-point allocation, the model bridges data gaps, enabling researchers to infer supply-chain structures and assess system behavior under diverse scenarios. The strong correspondence between ATP3's modeled flows and known pipeline configurations reinforces the reliability of its accuracy. Furthermore, the integration of fine-resolution emissions calculations showed that such models can extend beyond flow assignment to address environmental performance, opening avenues for combined energy systems and LCA research.

Energy policy perspective. The results yielded actionable insights for state-level energy sourcing and climate strategies. The mapped dominance of out-of-state gas supply to California's power sector underscores the strategic importance of interstate infrastructure and cooperation. By linking specific supply basins to individual power plants and quantifying their associated carbon intensities, ATP3's offers a data-driven foundation for targeted GHG reduction policies. For example, identifying import routes with disproportionately high carbon footprints can inform regulatory incentives for cleaner sourcing or for methane-leak mitigation in specific supply chains.

Industry perspective. For energy companies and utilities, ATP3 offers both supply-chain transparency and emissions accountability. By tracing plant-level gas origins and the associated carbon impacts, utilities can strengthen risk management, improve procurement strategies, and meet emerging sustainability reporting requirements. Through alternative sourcing or investment in certified low-carbon gas, high-emission segments can be identified for mitigation. Scenario analysis further supports infrastructure and contracting decisions, allowing operators to evaluate how supply reallocation may reshape flows, costs, and emissions profiles—aligning operational planning with the industry's growing emphasis on transparent, low-carbon supply networks.

Limitations. While the case study validated the methodological soundness and practical relevance of ATP3, several limitations should be noted. First, the accuracy of inferred flows is limited

by the granularity and reliability of available infrastructure and supply–demand datasets; where data are incomplete or outdated, modeled linkages may omit small-scale or recently commissioned assets. Second, the current formulation assumes steady-state operation and does not explicitly incorporate transient flow dynamics or seasonal demand variability, which could affect the applicability for short-term operational forecasting.

Overall, the California application illustrates the ability of ATP3 to connect theoretical modeling with tangible, real-world insights, thereby resolving complex gas supply patterns, validating them against empirical infrastructure data, and quantifying their environmental implications. By combining methodological rigor with decision-relevant outputs, ATP3 provides a versatile platform for scientific inquiry, policy design, and strategic industry planning.

5. Conclusions

This study introduces ATP3, a unified framework for natural gas transportation modeling that integrates graph-based network construction, EWM-based supply–demand allocation, and an MCF formulation with scenario-driven computational simplification. Designed to bridge the gap between granular infrastructure tracing and high-level energy system planning, ATP3 performs three complementary functions: ① the high-resolution tracing of energy flows and carbon intensities from source to end-use, ② cost- and emissions-aware network optimization, and ③ scenario simulation for exploring alternative decarbonization pathways.

Application to California's power generation sector demonstrated the ability of the model to accurately construct present-day supply linkages and align them with known interstate transmission infrastructure. Using the EWM-based 40% baseline allocation, ATP3 determined that approximately 90.53% of the gas consumed for electricity generation is imported, with Oregon, Nevada, and Arizona contributing 27.43%, 14.44%, and 48.66%, respectively. Plant-level results showed that Arizona is the single largest external supplier (1034.51 MMscf·d^{-1}), while in-state production accounts for only 201.37 MMscf·d^{-1} (9.47%). Source-specific allocation patterns revealed that Arizona and California each direct nearly half of their delivered gas to the power sector, in contrast to Oregon and Nevada, whose imports predominantly serve other end-uses. From a network-efficiency perspective, EWM reduced transportation costs by 2.76% relative to uniform allocation by favoring geographically proximate supply sources.

Looking ahead, the extensibility of ATP3 will allow it to remain a cutting-edge tool for energy transition analysis. Future work will focus on:

(1) Incorporating temporal and probabilistic dynamics. Expanding ATP3 to capture the temporal evolution in flows through integration with demand forecasting, probabilistic modeling, and power plant dispatch data, enabling analyses of phased technology deployment and transitional pathways through 2030, 2045, and beyond.

(2) Expanding geographic and sectoral scope. Extending ATP3 to encompass the full North American network and to incorporate end-user data across all demand sectors, thereby enabling comprehensive pathway tracing and transmission pattern analysis at both the geographic and sectoral scales.

(3) Integrating full life-cycle carbon and methane emissions modeling. Coupling the spatial and flow allocation outputs of ATP3 with detailed LCA databases, compressor station emission factors, transport-related carbon intensity metrics, and satellite-based methane observations, to support end-to-end GHG assessments and climate impact evaluations for each supply pathway.

Such enhancements will further strengthen ATP3's role in the modeling of decarbonized energy systems. As global climate efforts intensify, ATP3 is poised to become a pivotal decision-support tool for both researchers and practitioners. Its ability to unify traceability, optimization, and scenario analysis in one platform empowers stakeholders to navigate the complexity of energy transitions with confidence. In summary, ATP3 provides a robust foundation for devising and evaluating pathways to a sustainable, net-zero carbon future and will continue to guide informed decision-making on the journey toward decarbonization.

CRediT authorship contribution statement

Zemin Eitan Liu: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Diego Moya:** Writing – review & editing, Validation, Investigation. **Zhenlin Chen:** Writing – review & editing, Data curation. **Wennan Long:** Writing – review & editing, Conceptualization. **Liang Jing:** Validation, Funding acquisition, Data curation. **Bo Ren:** Writing – review & editing, Funding acquisition. **Haoyu Tang:** Investigation, Data curation. **Muhammad Y. Jabbar:** Writing – review & editing, Resources. **Farah Ramadan:** Writing – review & editing, Resources. **James Littlefield:** Writing – review & editing, Conceptualization. **Mohammad S. Masnadi:** Writing – review & editing, Validation, Supervision, Funding acquisition.

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.

Acknowledgments

Zemin Eitan Liu and Wennan Long would like to thank Aramco America Michigan office for financial support. We also wish to remember James Littlefield, whose insight and encouragement profoundly shaped this project; this work is dedicated to his memory.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eng.2025.11.005.

References

[1] Littlefield J, Rai S, Skone TJ. Life cycle GHG perspective on US natural gas delivery pathways. Environ Sci Technol 2022;56(22):16033–42.
[2] Liu ZE, Long W, Chen Z, Littlefield J, Jing L, Ren B, et al. A novel optimization framework for natural gas transportation pipeline networks based on deep reinforcement learning. Energy AI 2024;18:100434.
[3] Saad DM, Sodwatana M, Sherwin ED, Brandt AR. Energy storage in combined gas–electric energy transitions models: the case of California. Appl Energy 2025;385:125480.
[4] Lyu M, Zhang C, Bao X, Song J, Liu Z. Effects of the substitution rate of natural gas on the combustion and emission characteristics in a dual-fuel engine under full load. Adv Mech Eng 2017;9(12):1–8.
[5] Burns D, Grubert E. Attribution of production-stage methane emissions to assess spatial variability in the climate intensity of US natural gas consumption. Environ Res Lett 2021;16(4):044059.
[6] Lochran S. GNOME: a dynamic dispatch and investment optimisation model of the European natural gas network and its suppliers. Oper Res Forum 2021;2:67.
[7] Guo Z, Liu Z, Shuai S. Evolution and future development of vehicle fuel specification in China. Report. Warrendale: SAE Technical Paper; 2021.
[8] Welder L, Ryberg DS, Kotzur L, Grube T, Robinius M, Stolten D. Spatio-temporal optimization of a future energy system for power-to-hydrogen applications in Germany. Energy 2018;158:1130–49.
[9] Welder L, Stenzel P, Ebersbach N, Markewitz P, Robinius M, Emonts B, et al. Design and evaluation of hydrogen electricity reconversion pathways in

national energy systems using spatially and temporally resolved energy system optimization. Int J Hydrogen Energy 2019;44(19):9594–607.
[10] Samsatli S, Staffell I, Samsatli NJ. Optimal design and operation of integrated wind–hydrogen–electricity networks for decarbonising the domestic transport sector in great Britain. Int J Hydrogen Energy 2016;41(1):447–75.
[11] Moreno-Benito M, Agnolucci P, Papageorgiou LG. Towards a sustainable hydrogen economy: optimisation-based framework for hydrogen infrastructure development. Comput Chem Eng 2017;102:110–27.
[12] Roman-White SA, Mallikarjuna Prasanna D, McCullagh A, Ravikumar AP, Allen DT, Chivukula K, et al. Gas pathing: improved greenhouse gas emission estimates of liquefied natural gas exports through enhanced supply chain resolution. ACS Sustain Chem Eng 2024;12(46):16956–66.
[13] Gabriel SA, Kydes AS, Whitman P. The national energy modeling system: a large-scale energy-economic equilibrium model. Oper Res 2001;49(1):14–25.
[14] Müller-Kirchenbauer J, Ragwitz M, Kneiske T, Klaassen B, Mielich T, Herrmann U. A network modeling systematics for transition paths toward climate neutral gas networks—NeMoSys. Energy Technol 2025;13(2):2300977.
[15] Di Lullo G, Oni AO, Gemechu E, Kumar A. Developing a greenhouse gas life cycle assessment framework for natural gas transmission pipelines. J Nat Gas Sci Eng 2020;75:103136.
[16] 2022 California gas report. Report. Los Angeles: SoCalGas; 2022.
[17] Seebregts AJ, Goldstein GA, Smekens K. Energy/environmental modeling with the MARKAL family of models. In: Operations Research Proceedings 2001; 2001 Sep 3–5; Duisburg, Germany. Springer; 2001. p. 75–82.
[18] Calvin K, Patel P, Clarke L, Asrar G, Bond-Lamberty B, Cui RY, et al. GCAM v5. 1: representing the linkages between energy, water, land, climate, and economic systems. Geosci Model Dev 2019;12(2):677–98.
[19] Capros P, Kannavou M, Evangelopoulou S, Petropoulos A, Siskos P, Tasios N, et al. Outlook of the EU energy system up to 2050: the case of scenarios prepared for European Commission's "clean energy for all Europeans" package using the PRIMES model. Energ Strat Rev 2018;22:255–63.
[20] Sakellaris K, Canton J, Zafeiratou E, Fournié L. METIS—an energy modelling tool to support transparent policy making. Energ Strat Rev 2018;22:127–35.
[21] Mintz M, Elgowainy A, Gillette J, Paster M, Ringer M, Brown D, et al. HDSAM 2.0: expanded capabilities, enhanced results in hydrogen delivery modeling. In: Proceedings of The NHA Annual Hydrogen Conference 2008; 2008 Apr 1; Sacramento, CA, USA. National Hydrogen Association; 2008.
[22] Ulvestad M, Overland I. Natural gas and CO_2 price variation: impact on the relative cost-efficiency of LNG and pipelines. Int J Environ Stud 2012;69 (3):407–26.
[23] Greene S, Jia H, Rubio-Domingo G. Well-to-tank carbon emissions from crude oil maritime transportation. Transp Res Part D Transp Environ 2020;88:102587.
[24] Zubair M, Chen S, Ma Y, Hu X. A systematic review on carbon dioxide (CO_2) emission measurement methods under PRISMA guidelines: transportation sustainability and development programs. Sustainability 2023;15(6):4817.
[25] Winarno E, Hadikurniawati W, Rosso RN. Location based service for presence system using haversine method. In: Proceedings of 2017 International Conference on Innovative and Creative Information Technology (ICITech); 2017 Nov 2–4; Salatiga, Indonesia. IEEE; 2017. p. 1–4.
[26] Wang G, Cheng Q, Zhao W, Liao Q, Zhang H. Review on the transport capacity management of oil and gas pipeline network: challenges and opportunities of future pipeline transport. Energ Strat Rev 2022;43:100933.
[27] Boyles SD, Lownes NE, Unnikrishnan A. Transportation network analysis. Volume I: static and dynamic traffic assignment 2020. 2025. arXiv:2502.05182.
[28] Ahuja RK, Magnanti TL, Orlin JB. Network flows. Kent: Prentice Hall; 1988.
[29] Dijkstra EW. A note on two problems in connexion with graphs. In: Apt KR, Hoare T, editors. Edsger wybe dijkstra: his life, work, and legacy. New York City: ACM; 2022. p. 287–90.
[30] Chvátal V. Linear programming. London: Macmillan; 1983.
[31] Background information on the structures and functioning of the natural gas production and delivery systems that serve California and the US 2006. Report. Sacramento: California Energy Commission; San Francisco: California Public Utilities Commission; 2006.
[32] Gill L, Gutierrez A, Weeks T. 2021 SB 100 joint agency report: Achieving 100 percent clean electricity in California, an initial assessment. Report. Sacramento: California Energy Commission; 2021.
[33] Power plants. HIFLD; 2022.
[34] Natural gas processing plant survey 2017. Report. Washington, DC: US Energy Information Administration; 2017.
[35] El-Houjeiri HM, Brandt AR, Duffy JE. Open-source LCA tool for estimating greenhouse gas emissions from crude oil production using field characteristics. Environ Sci Technol 2013;47(11):5998–6006.
[36] Chen Z, Zhong R, Long W, Tang H, Wang A, Liu Z, et al. Advancing oil and gas emissions assessment through large language model data extraction. Energy AI 2025;20:100481.
[37] Chen Z, Zhong R, Long W, Tang H, Wang A, Liu Z, et al. AI-driven environmental data extraction for energy sector assessment. In: Proceedings of SPE Annual Technical Conference and Exhibition; 2024 Sep 23–25; New Orleans, LA, USA. OnePetro; 2024.
[38] Innovative data energy applications 2015. Report. Golden: National Renewable Energy Laboratory; 2015.
[39] Masnadi MS, Long W, Jing L, Liu ZE, Jabbar MY, Chen Z, et al. Global carbon intensity of LNG value chain amid intensifying energy security and climate trade-offs. Res Square 2025.
[40] Natural gas and California [Internet]. San Francisco: California Public Utilities Commission (CPUC); [cited 2025 Aug 12]. Available from: https://www.cpuc.ca.gov/industries-and-topics/natural-gas/natural-gas-and-california.