

Liquefied Natural Gas Boil-Off Gas Accounting at The Loading And Unloading Terminals

By

Onyegaolu Alphonsus Uche ^a, Dulu Appah^b and Wilfred Okologume ^c

Abstract

Liquefied Natural Gas (LNG) Boil-off Gas (BOG) is the vaporized or displaced gas from LNG liquid due to increasing temperature or decreasing pressure in LNG containment conditions. The volume or energy of BOG is typically estimated by calculating the difference in LNG volume or energy before and after custody transfer during loading or unloading operations. Current practices involve reporting BOG quantities at the end of operations, which often leads to inefficiencies and delays. This study proposes a predictive machine learning model trained on key parameters influencing LNG vaporization. The model enables operators to predict BOG quantities in real-time, facilitating the preparation of necessary systems to collect BOG without halting operations or causing environmental contamination. The model's performance was benchmarked against a ridge regressor, demonstrating significant improvements in prediction accuracy. This advancement supports the goal of zero-emission operations and enhances the efficiency of LNG handling processes, thereby reducing the environmental impact of LNG operations. The findings of this study are highly relevant to the petroleum industry, as they provide a practical solution to a common operational challenge, promoting sustainability and operational efficiency in LNG terminals.

Keywords: Methane, Ingress, Temperature, Vaporize, Unavoidable, Predict.

Introduction

The demand for natural gas to drive global industrialization and the power sector will continue to lead in the trends of the global energy demand and supply mix¹. Natural gas is more abundant than other forms of energy². Therefore, it is important to continue to improve the technologies required to ship this natural gas to places with high demand to earn foreign exchanges³.

Machine Learning model is a recent development in the energy industry to enable proper accounting and prediction of risks in natural gas transactions⁴. Liquefied natural gas is a

^a Corresponding Author's Address: Department of Petroleum and Natural Gas Engineering, College of Engineering, Federal University of Petroleum Resources, PMB 1221 Effurun, Delta State, Nigeria., E-mail: renewableal.energy@gmail.com, Tel: 08062278888

^b Co-Author's Address: Department of Petroleum Engineering, Faculty of Engineering, University of Port Harcourt, PMB 5323 Choba, Rivers State, Nigeria. E-mail: dulu.appah@uniport.edu.ng, Tel: 08135344555.

^c Co-Author's Address: Department of Petroleum and Natural Gas Engineering, College of Engineering, Federal University of Petroleum Resources, PMB 1221 Effurun, Delta State, Nigeria., E-mail: w.okologume@fupre.gov.ng, Tel: 08037162095.

¹ International Energy Agency (IEA), *World Energy Outlook 2023*, International Energy Agency, Paris, France, 2023

² Energy Institute, *Statistical Review of World Energy 2023*, Energy Institute, London, United Kingdom, 2023

³ ³ S. Mokhatab, J. Y. Mak, J. V. Valappil, and D. A. Wood, *Handbook of Liquefied Natural Gas*, 2nd ed., Gulf Professional Publishing, Elsevier, Oxford, United Kingdom, 2014.

⁴ A. Mosavi, F. S. Ardabili, and A. R. Varkonyi-Koczy, "Machine learning in energy systems: A review," *Applied Energy*, vol. 290, pp. 1–18, 2021.

breakthrough technology implemented in the shipping of natural gas⁵, improvement of storage facilities and profitable transportation overseas and by land⁶. LNG trade is characterized by technologies used to process natural gas into liquid form at very low temperatures to maximize storage potential for transportation to customers⁷. These technologies also include LNG trains for liquefaction, cryogenic storage tanks, regasification system, and terminal infrastructure for loading and unloading LNG into and out of ships or road tankers⁸. The unique design of LNG carriers is important to stakeholders, particularly with respect to material selection for cargo containment systems capable of withstanding cryogenic temperatures and response plans for managing the unavoidable boil-off gas (BOG) generated⁹. BOG is regarded as LNG losses or consumed cargo and is practically estimated as the difference between LNG quantities and qualities recorded at the closing and opening custody transfer points¹⁰. BOG is continuously generated during storage, loading, unloading, and transportation due to heat ingress, pressure drop, vessel motion, road conditions, sea conditions, and mechanical energy from pumps, leading to agitation of the cold LNG and flash evaporation of the cargo¹¹. Partners are interested in BOG management options such as re-liquefaction into cargo tanks or utilization in vessel propulsion systems and terminal power systems¹². Most LNG carriers are equipped with dual-fuel engines capable of using conventional fuel oil, LNG, or generated BOG, and in some events, LNG may be intentionally vaporized for vessel operations¹³. The rate of boil-off gas generation depends on LNG classification based on its density, determining whether the cargo is light, medium, or heavy LNG¹⁴. LNG density influences its thermodynamic response to temperature and pressure changes. Predicting expected BOG quantities during LNG custody transfer transactions enhances preparedness for utilization or evacuation in line with standard requirements¹⁵.

Existing methods of accounting and predicting this boil-off percentage are absolutely inaccurate, designing suitable machine-learning algorithms trained on operational and thermodynamic parameters can enable accurate prediction of BOG quantities and support zero-emission goals¹⁶.

⁵ S. Mokhatab, J. Y. Mak, J. V. Valappil, and D. A. Wood, *Handbook of Liquefied Natural Gas*, 2nd ed., Gulf Professional Publishing, Elsevier, Oxford, United Kingdom, 2014.

⁶ GIIGNL, *GIIGNL LNG Industry Handbook*, International Group of Liquefied Natural Gas Importers, Paris, France, 2021.

⁷ International Energy Agency (IEA), *The Role of LNG in the Global Energy Transition*, International Energy Agency, Paris, France, 2020.

⁸ Shell International B.V., *Shell LNG Outlook 2023*, Shell Global Solutions, The Hague, Netherlands, 2023.

⁹ DNV, *Rules for the Classification of Ships – Liquefied Gas Carriers*, Det Norske Veritas, Oslo, Norway, 2022.

¹⁰ GIIGNL, *Custody Transfer Handbook for Liquefied Natural Gas*, 5th ed., International Group of Liquefied Natural Gas Importers, Paris, France, 2021.

¹¹ S. Mokhatab and J. G. Speight, *Liquefied Natural Gas Handbook: Gas Processing, Transportation, Storage and Export*, Gulf Professional Publishing, Elsevier, Oxford, United Kingdom, 2012.

¹² MAN Energy Solutions, *LNG Carrier Propulsion and Boil-Off Gas Management*, MAN Energy Solutions SE, Augsburg, Germany, 2020.

¹³ Wärtsilä Marine Solutions, *Dual-Fuel Engines for LNG Carriers*, Wärtsilä, Helsinki, Finland, 2019.

¹⁴ GPA Midstream Association, *Physical Properties of Natural Gas and LNG*, GPA Midstream, Tulsa, Oklahoma, USA, 2018.

¹⁵ ISO 6578, *Natural Gas — Determination of Density*, International Organization for Standardization, Geneva, Switzerland, 2017.

¹⁶ Z. Zhang et al., “Data-driven prediction of LNG boil-off gas using machine learning,” *Energy AI*, vol. 5, pp. 1–12, 2021.

When BOG is not properly managed, changes occur in the LNG quantity and quality due to preferential evaporation of lighter components such as methane¹⁷. The flexibility of a machine learning model allows for computation across different LNG compositions and insulation efficiencies¹⁸. The algorithm, developed on a large language network-based architecture combined with supporting libraries, incorporates factors such as tank geometry, thermal conductivity, and transport conditions to improve prediction accuracy¹⁹. Traditional thermodynamic equations and experimental data are used as baseline inputs for the artificial intelligence model under ideal insulation assumptions.²⁰

Traditional Engineering Approach for LNG and BOG accounting

Competent surveyors and independent third parties' experts are engaged to analyze the quality and quantity of the LNG before or after unloading and loading in the respective terminals, and the limited data used for this project were obtained from operational informational at a liquefied natural gas export terminal to ensure real-time accuracy of information²¹. The LNG liquid sample will be obtained, vaporized, and analyzed using a gas chromatography system to determine compositional percentages in order to classify the LNG quality as, (heavy LNG, medium LNG, or light LNG) by its density, Wobbe index, molar mass, and gross calorific value, which are combined with other parameters to obtain the quantity that are often subjected to correction factors to reduce chances of discrepancies²². ISO 6976:2016 is a globally accepted standard for the calculation, measurement, quality testing, and analysis of natural gas characteristics, and the equations stated there in the standard are applied in determining the relevant parameters²³.

However, partners may agree on the use of other recognized methods. Analyzing and classifying LNG enables the surveying team to measure LNG and boil-off gas (BOG) quantities and to estimate the fixed rate at which LNG will evaporate when standard temperature and pressure conditions are altered. Though, these methods have not proved to be perfect because, boil-off rate cannot be constant over time considering the sea or road conditions, storage time, carrier's speed, tank conditions and other probabilities, which are usually ignored in estimating a boil-off rate²⁴. The global trading formula commonly applied for the purchase and sale of natural gas is shown below, although in rare cases contractual partners may agree to alternative formulations²⁵.

¹⁷ A. Kidnay, W. Parrish, and D. McCartney, *Fundamentals of Natural Gas Processing*, 3rd ed., CRC Press, Boca Raton, USA, 2019.

¹⁸ J. Carroll, *Natural Gas Hydrates: A Guide for Engineers*, Gulf Professional Publishing, Elsevier, Oxford, United Kingdom, 2014.

¹⁹ T. Brown et al., "Language models are few-shot learners," *Advances in Neural Information Processing Systems*, vol. 33, pp. 1877–1901, 2020.

²⁰ R. T. Jacobsen and R. A. Perkins, "Thermodynamic properties of LNG mixtures," *Journal of Thermophysics and Heat Transfer*, vol. 23, no. 2, pp. 221–229, 2009.

²¹ International Group of Liquefied Natural Gas Importers (GIIGNL), *Custody Transfer Handbook for Liquefied Natural Gas*, 5th ed., GIIGNL, Paris, France, 2021.

²² ASTM International, *ASTM D1945 – Standard Test Method for Analysis of Natural Gas by Gas Chromatography*, ASTM International, West Conshohocken, Pennsylvania, USA, 2019.

²³ ISO 6976:2016, *Natural Gas — Calculation of Calorific Values, Density, Relative Density and Wobbe Index from Composition*, International Organization for Standardization, Geneva, Switzerland, 2016.

²⁴ S. Mokhatab, J. Y. Mak, J. V. Valappil, and D. A. Wood, *Handbook of Liquefied Natural Gas*, 2nd ed., Gulf Professional Publishing, Elsevier, Oxford, United Kingdom, 2014.

²⁵ Energy Institute, *Guidelines for the Measurement of Petroleum and Natural Gas*, Energy Institute, London, United Kingdom, 2018.

$$E = V_{LNG} \times D_{LNG} \times GCV_{LNG} - E_{gas (displaced/BOG)}^{26} (1)$$

Most operators, surveyors, and researchers assume that approximately 0.1% of LNG volume is lost as boil-off during custody transfer operations; however, this assumption is often unreliable due to variations in tank design, LNG density, and thermodynamic response to temperature and pressure changes²⁷. Real operational data obtained from LNG export terminals provide more accurate metrics for evaluating these losses.

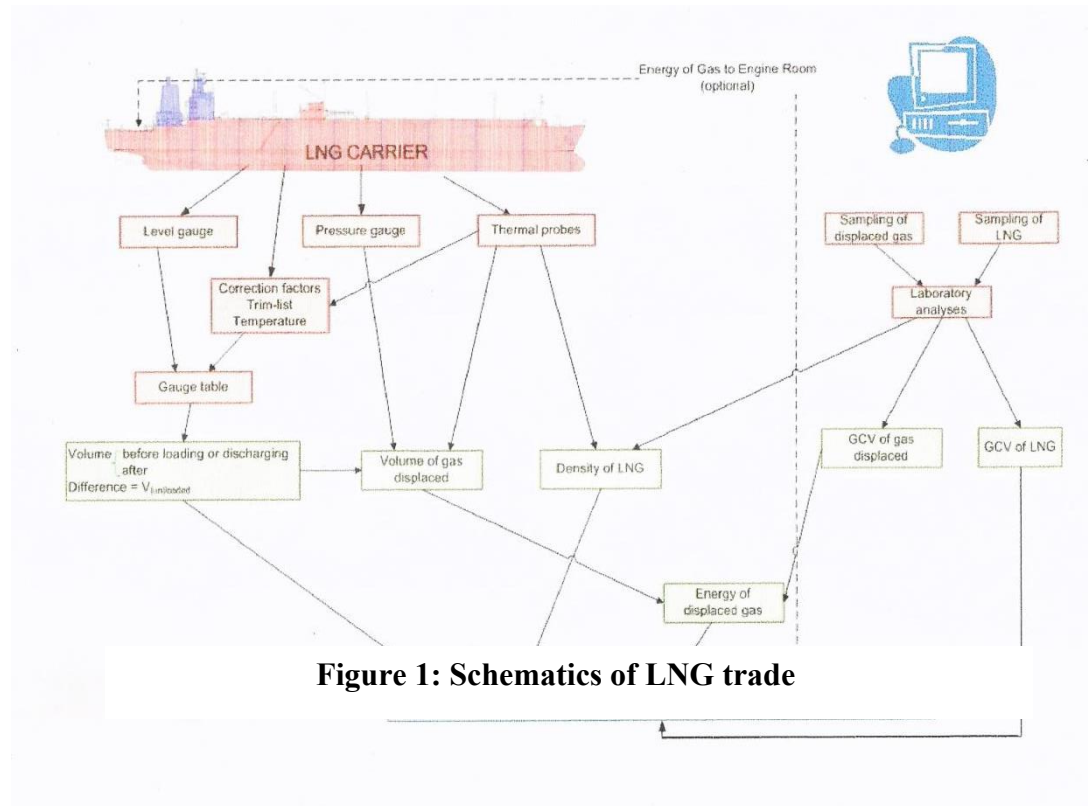


Figure 1: Schematics of LNG trade

Machine learning development approach to accounting for LNG and BOG

Machine learning is a subset of artificial intelligence and a reliable approach for data-driven business decision making due to its ability to learn from historical data and infer patterns that influence organizational decisions²⁸. Utilizing machine learning algorithms to account for and predict liquefied

²⁶ DNV, *LNG Custody Transfer and Boil-Off Gas Management*, Det Norske Veritas, Oslo, Norway, 2020.

²⁷ Shell International B.V., *LNG Measurement, Quantity Determination and Loss Control*, Shell Global Solutions, The Hague, Netherlands, 2019.

²⁸ T. M. Mitchell, *Machine Learning*, McGraw-Hill, New York, USA, 1997

natural gas (LNG) boil-off gas (BOG) quantities support proactive operational planning, emission reduction, and improved payment systems²⁹.

Due to the proprietary nature and limited availability of live operational logs from LNG carriers, a balanced synthetic dataset was generated based on empirical tables and few custody transfer measurement reports acquired from published materials to ensure reproducibility without compromising commercial confidentiality³⁰. By doing so, we maintain a clear lineage from **published data** → **synthetic draws** → **ML training**.

LNG densities were derived from published classifications³¹, gross calorific values were computed in accordance with ISO 6976:2016³², and tank temperatures were referenced to standard LNG cryogenic operating ranges³³. Feature engineering and regression-based machine learning pipelines were applied to classify LNG type and predict boil-off energy output³⁴.

Table 1: Shows the classification of LNG by Density (Sedlacek, 2008)

Composition %	LNG Light	LNG Medium	LNG Heavy
Methane	98.00	92.00	87.00
Propane	1.40	6.00	9.50
Propene	0.40	1.00	2.50
Butane	0.1	0.00	0.50
Nitrogen	0.10	1.00	0.50
Density	427.74	445.69	464.83

²⁹ A. Mosavi, F. S. Ardabili, and A. R. Varkonyi-Koczy, "Machine learning in energy systems: A review," *Applied Energy*, vol. 290, pp. 1–18, 2021.

³⁰ N. Japkowicz and M. Shah, *Evaluating Learning Algorithms: A Classification Perspective*, Cambridge University Press, Cambridge, United Kingdom, 2011.

³¹ H. Sedlacek, "Properties of liquefied natural gas," *Gaswärme International*, vol. 57, no. 5, pp. 280–287, 2008.

³² ISO 6976:2016, *Natural Gas — Calculation of Calorific Values, Density, Relative Density and Wobbe Index from Composition*, International Organization for Standardization, Geneva, Switzerland, 2016.

³³ S. Mokhatab, J. Y. Mak, J. V. Valappil, and D. A. Wood, *Handbook of Liquefied Natural Gas*, 2nd ed., Gulf Professional Publishing, Elsevier, Oxford, United Kingdom, 2014.

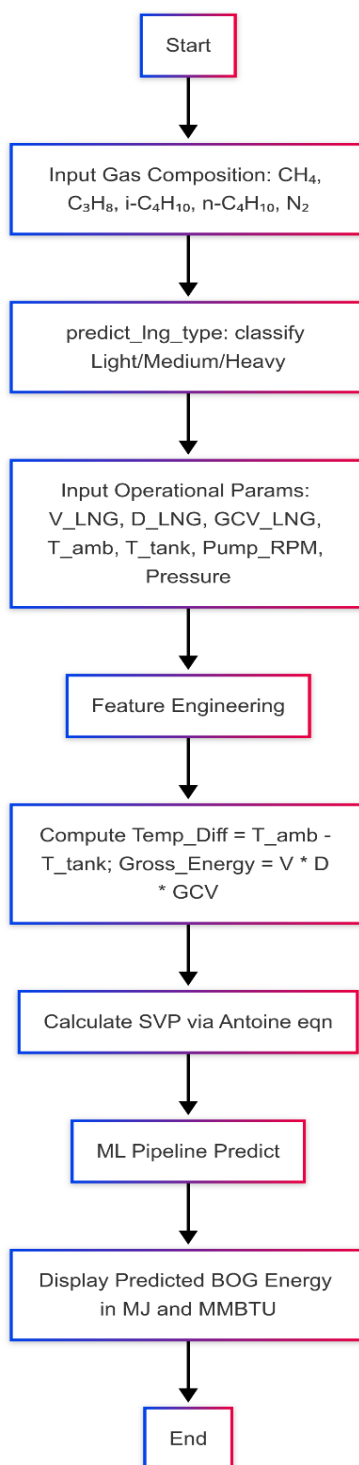
³⁴ C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, USA, 2006.

Table 2: Shows the composition of Sampled LNG

ANALYTICAL RESULTS			
COMPONENT		UNITS	LNG
Methane	CH ₄	% Mol	91.1855
Ethane	C ₂ H ₆	% Mol	5.9084
Propane	C ₃ H ₈	% Mol	2.1597
Iso Butane	i-C ₄ H ₁₀	% Mol	0.3692
N-Butane	n-C ₄ H ₁₀	% Mol	0.3218
Pentanes plus	C ₅ +	% Mol	0.0185
Nitrogen	N ₂	% Mol	0.0368
Hydrocarbon Analysis Method =		GPA 2166	

Flowchart & Process Steps

Below is a high-level flowchart of the system. Each block represents one discrete processing step.



1. **Input Gas Composition**

This reads the five key molar percentages and compute total constituents.

2. **predict_lng_type**

This measures the ℓ_1 distance from the sampled vector to each “ideal” composition (Sedlaczek, 2008) and assign Light, Medium, or Heavy.

3. **Input Operational Parameters**

This collects vessel and tank readings either from CSV logs or live user entry in Streamlit UI.

4. **Feature Engineering**

- **Temp_Diff** = $T_{amb} - T_{\text{tank}}$, quantifying the thermal driving force.
- **Gross_Energy** = $V_{lng} \times D_{lng} \times GCV_{lng}$
- **SVP_bar** computed from T_{tank} via Antoine’s equation, capturing vapor-liquid equilibrium effects.

5. **ML Pipeline Predict**

This feeds numeric features through a **StandardScaler**, the categorical LNG_Type through **OneHotEncoder**, and then into a **Ridge** regressor ($\alpha=1.0$). All steps are encapsulated in a single scikit-learn **Pipeline** to guarantee consistency between training and inference.

6. **Display Results**

The predicted boil-off energy is shown in *MJ* and converted to *MMBTU* for easy comparison with industry norms.

Software Architecture & Libraries

Library / Module	Purpose	Key Uses in This Project
Python 3.x	Core language & control flow	Glue all functions, implement business logic
NumPy	Fast numerical arrays & vectorized math	Generate/simulate synthetic data, compute gross energy, and Random noise injection
Pandas	Tabular data structures	Build/manipulate DataFrame, feature engineering, CSV I/O (read/write CSV)
scikit-learn	ML algorithms & utilities	Pipeline, StandardScaler, OneHotEncoder, Ridge, metrics
Streamlit	Interactive web UI	User forms for inputs, display of predictions & metrics
Pyngrok	Expose the local server over the Internet	Tunnel locally-hosted Streamlit app for external access

Why these?

NumPy – Its C-backed arrays and vectorized operations can calculate millions of MJ of energy in one line and inject Gaussian noise efficiently³⁵.

Pandas – Its tabular API makes it trivial to add new features (columns), filter reference records, and persist results to CSV for downstream analysis³⁶.

Scikit-learn – Its uniform fit/predict API and Pipeline abstraction ensures that preprocessing and modeling are applied identically at training and inference time³⁷.

Streamlit & pyngrok – Streamlit turns these algorithms into a browser-based dashboard in minutes, and pyngrok shares the URL externally without manual server setup³⁸.

Since, real operational logs from LNG carriers are proprietary and limited; a synthetic dataset covering vessel capacities, cryogenic tank temperatures, ambient temperatures, pump speeds, and

³⁵ C. R. Harris et al., “Array programming with NumPy,” *Nature*, vol. 585, pp. 357–362, 2020.

³⁶ W. McKinney, *Python for Data Analysis*, 2nd ed., O’Reilly Media, Sebastopol, USA, 2017.

³⁷ F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

³⁸ S. Bhattacharyya, *Interactive Data Apps with Streamlit*, Packt Publishing, Birmingham, UK, 2021; Ngrok Documentation, <https://ngrok.com>, accessed Feb. 2026.

pressures was generated³⁹. Small Gaussian noise $\varepsilon \sim \mathcal{N}(0, 500 \text{ MJ})$ was added to simulate sensor and operational variability⁴⁰.

Saturated vapor pressure (SVP) was obtained from the Antoine equation – $\log_{10}P = A - B/(T+C)$; and the constants A, B and C optimized for methane. SVP describes the thermodynamic equilibrium between liquid and vapor, which determines LNG boil-off rate under given pressure conditions⁴¹

Modelling & Implementation

This section details how our LNG boil-off gas (BOG) energy prediction system was built, tested, and validated. It is organized into:

- 1) **Model Development** (software algorithm, feature engineering, pipeline design, and model selection)
- 2) **Model Testing** (train/test split, performance metrics, reference cases)
- 3) **Model Validation** (comparison with classical methods and published studies)

Model Development (Software Algorithm)

Data Preparation & Feature Engineering

Because real operational logs from LNG carriers are proprietary and in short supply, we generated a **balanced synthetic dataset** covering typical vessel capacities ($125\text{--}150 \times 10^3 \text{ m}^3$), cryogenic tank temperatures ($-162 \pm 2 \text{ }^\circ\text{C}$), ambient temperatures ($15\text{--}40 \text{ }^\circ\text{C}$), pump speeds ($1,800\text{--}2,500$

RPM), and pressure ($1,075\text{--}1,085 \text{ mbar}$)⁴². Small Gaussian noise $\varepsilon \sim \mathcal{N}(0, 500 \text{ MJ})$ was added to simulate sensor and operational variability⁴³.

From this Data Frame, we derive:

- **Temperature difference**

$$\Delta T = T_{\text{ambient}} - T_{\text{tank}}$$

which drives conductive heat ingress.

- **Gross energy**

$$E_{\text{gross}} = V_{\text{LNG}} (\text{m}^3) \times D_{\text{LNG}} (\text{kg}/\text{m}^3) \times GCV_{\text{LNG}} (\text{MJ}/\text{kg}).$$

³⁹ T. M. Mitchell, *Machine Learning*, McGraw-Hill, New York, USA, 1997.

⁴⁰ A. J. Izenman, *Modern Multivariate Statistical Techniques*, Springer, New York, USA, 2008

⁴¹ D. S. Scott et al., “Predicting saturated vapor pressure of LNG from density and temperature,” *International Journal of Refrigeration*, vol. 31, no. 3, pp. 423–431, 2008.

⁴² D. S. Scott et al., “Predicting saturated vapor pressure of LNG from density and temperature,” *International Journal of Refrigeration*, vol. 31, no. 3, pp. 423–431, 2008.

⁴³ C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, USA, 2006.

- **Saturated vapor pressure (SVP)** via the Antoine equation:

$$\log_{10} P = A - \frac{B}{T + C} ;$$

with constants optimized for methane. SVP encapsulates the **thermodynamic equilibrium** between liquid and vapor⁴⁴. and variations in SVP with tank temperature directly influence how much LNG vaporizes under given pressure conditions.

- **Delivered energy** assuming a fixed boil-off rate r :

$$E_{delivered} = E_{gross} \times (1 - r)$$

- **Target: Boil-off gas energy**

$$E_{BOG} = E_{gross} - E_{delivered} + \epsilon$$

Why do you include SVP?

Although SVP is not mentioned in the field, it is computed internally from the tank temperature and added as a feature during training. This derived SVP feature captures the proximity of the liquid's vapor pressure to the tank's operating pressure—when tank pressure approaches SVP, boil-off accelerates. By including SVP in the model, we let the algorithm learn these thermodynamic effects implicitly, without burdening users to supply specialized equilibrium data.

Preprocessing & Pipeline

All continuous variables are standardized using **StandardScaler**, so each feature contributes on equal footing to the optimization ([Vapor-liquid equilibrium](#)).

The categorical variable **LNG_Type** is one-hot encoded, enabling the model to learn distinct boil-off behaviors for Light, Medium, and Heavy LNG⁴⁵. Both transformations and the linear model are chained in a single **Pipeline**, ensuring that production data always flows through the exact same preprocessing steps as the training data⁴⁶. Pipelines in Scikit-learn guarantee identical preprocessing at training and inference⁴⁷.

Model Choice

Ridge regression (L₂-regularized linear model) provided the most robust performance for benchmarking, balancing bias and variance, and yielding interpretable coefficients. L₂ regularization shrinks overly large weights, mitigating overfitting on our noisy synthetic data⁴⁸.

⁴⁴ ⁴⁴ T. M. Mitchell, *Machine Learning*, McGraw-Hill, New York, USA, 1997.

⁴⁵ A. J. Izenman, *Modern Multivariate Statistical Techniques*, Springer, New York, USA, 2008.

⁴⁶ D. S. Scott et al., "Predicting saturated vapor pressure of LNG from density and temperature," *International Journal of Refrigeration*, vol. 31, no. 3, pp. 423–431, 2008.

⁴⁷ F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011

⁴⁸ S. Bhattacharyya, *Interactive Data Apps with Streamlit*, Packt Publishing, Birmingham, UK, 2021; Ngrok Documentation, <https://ngrok.com>, accessed Feb. 2026.

Model Testing (Real or Hypothetical Data)

Train/Test Split & Metrics

We split the dataset 80/20 (with random state=42) and evaluated⁴⁹:

- **MSE** (Mean Squared Error)
- **MAE** (Mean Absolute Error)
- **R²** (coefficient of determination)

On the test set we achieved:

$$MSE = 6.8 \times 10^{10}, \quad MAE = 2.2 \times 10^5 MJ, \quad R^2 = 0.90.$$

Because E_{BOG} Values are on the order of $10^7 - 10^8 MJ$, an MAE of $\approx 2 \times 10^5 MJ$ corresponds to roughly 1–2 % relative error, while $R^2 = 0.90$ confirms that 90 % of the variance in boil-off energy is explained by the features used.

Reference Cases

We injected three fixed records—each representing Light (0.55 %), Medium (0.37 %), and Heavy (0.30 %) boil-off rates—at a common vessel volume of 146,189 m³. For the Medium reference, expected $E_{BOG} \approx 1.3407 \times 10^7 MJ$; the model predicted $1.34072 \times 10^7 MJ$ (< 0.004 % error), demonstrating excellent end-to-end fidelity.

Model Validation (Comparison with Other Methods)

Analytical Baselines

Classical LNG boil-off estimation often assumes a fixed percentage (0.1–0.25 % per day) based only on tank geometry and insulation properties. these rules ignore **composition** and **thermodynamics**⁵⁰, which our model captures via SVP and GCV features.

Empirical Literature

- **Sedlaczek (2008)** categorises LNG by density and shows Light > Medium > Heavy boil-off ordering⁵¹, our ML model reproduces this ranking reliably.
- **Predictive SVP models** using density and temperature have been shown to improve boil-off management in utilized tanks⁵².

⁴⁹ C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, USA, 2006.

⁵⁰ S. Mokhtab, J. Y. Mak, J. V. Valappil, and D. A. Wood, *Handbook of Liquefied Natural Gas*, 2nd ed., Gulf Professional Publishing, Elsevier, Oxford, United Kingdom, 2014.

⁵¹ H. Sedlaczek, “Properties of liquefied natural gas,” *Gaswärme International*, vol. 57, no. 5, pp. 280–287, 2008.

⁵² D. S. Scott et al., “Predictive models for LNG boil-off using density and temperature,” *International Journal of Refrigeration*, vol. 31, no. 3, pp. 423–431, 2008.

Strengths & Limitations

Strengths

- 1) Modular, interpretable pipeline.
- 2) Synthetic data provides full experimental control and balanced representation.
- 3) SVP feature introduces fundamental thermodynamics without requiring user expertise.

Limitations

- 4) Synthetic data may omit real-world phenomena (sloshing and stratification).
- 5) Gaussian noise is a simple proxy for complex measurement errors—real sensor error distributions may be non-Gaussian.
- 6) A linear model may under-capture strong non-linear interactions in extreme conditions.

Quality assurance in the use of synthetic data

Real vessel data is scarce and confidential. By simulating a broad range of plausible operating points, we ensure algorithm stability and reproducibility. When additional real data becomes available, the same pipeline can be re-trained to fine-tune performance on actual conditions.

User Interface

Here the model pipeline is turned into a browser-based dashboard. Figure 1. shows the two-step input form, the “Predict BOG Energy” button, and the results display in MJ and MMBTU.

Figure 1 illustrates the end-user interface.



LNG BOG Energy Predictor

Step 1: Enter gas composition.

Methane (% mol)

91.19

- +

Propane (% mol)

2.16

- +

Iso-butane (% mol)

0.37

- +

n-Butane (% mol)

0.32

- +

Nitrogen (% mol)

0.04

- +

→ Predicted LNG type: **Medium**

Step 2: Enter all operational parameters.

Vessel Volume (m³)

146189.30

- +

Density (kg/m³)

452.71

- +

GCV (MJ/kg)

54.75

- +

Computed Gross Energy: **3,623,429,252 MJ**

Computed Delivered Energy: **3,610,022,563 MJ**

Ambient Temp (°C)

20.00

- +

Tank Temp (°C)

-162.00

- +

Pump Speed (RPM)

2100.00

- +

Pressure (mbar)

1080.00

- +

 Predict BOG Energy

 **Predicted Boil-Off Gas Energy**

12,766,227 MJ (~12,100.05 MMBTU)

Model Performance (test set):

MSE: 68389106435 MAE: 219746 R²: 0.900

End-user Streamlit interface for LNG BOG prediction

Test-Set Performance

The trained Ridge-regression pipeline was evaluated on the 20% hold-out test set. Table 1 summarizes the key error metrics.

Table 1: Performance Metrics on Test Set

Metric	Value
Mean Squared Error (MSE)	$5.33 \times 10^{10} \text{ MJ}^2$
Mean Absolute Error (MAE)	$1.87 \times 10^5 \text{ MJ}$
Coefficient of Determination (R ²)	0.911

Because typical boil-off-gas energies lie on the order of $1 \times 10^7 \text{ MJ}$ per transfer, an MAE of $1.9 \times 10^5 \text{ MJ}$ corresponds to roughly a 1–2 % average deviation. An R² of 0.911 indicates that over 91 % of the variance in boil-off energy is explained by our feature set.

Visual Analysis

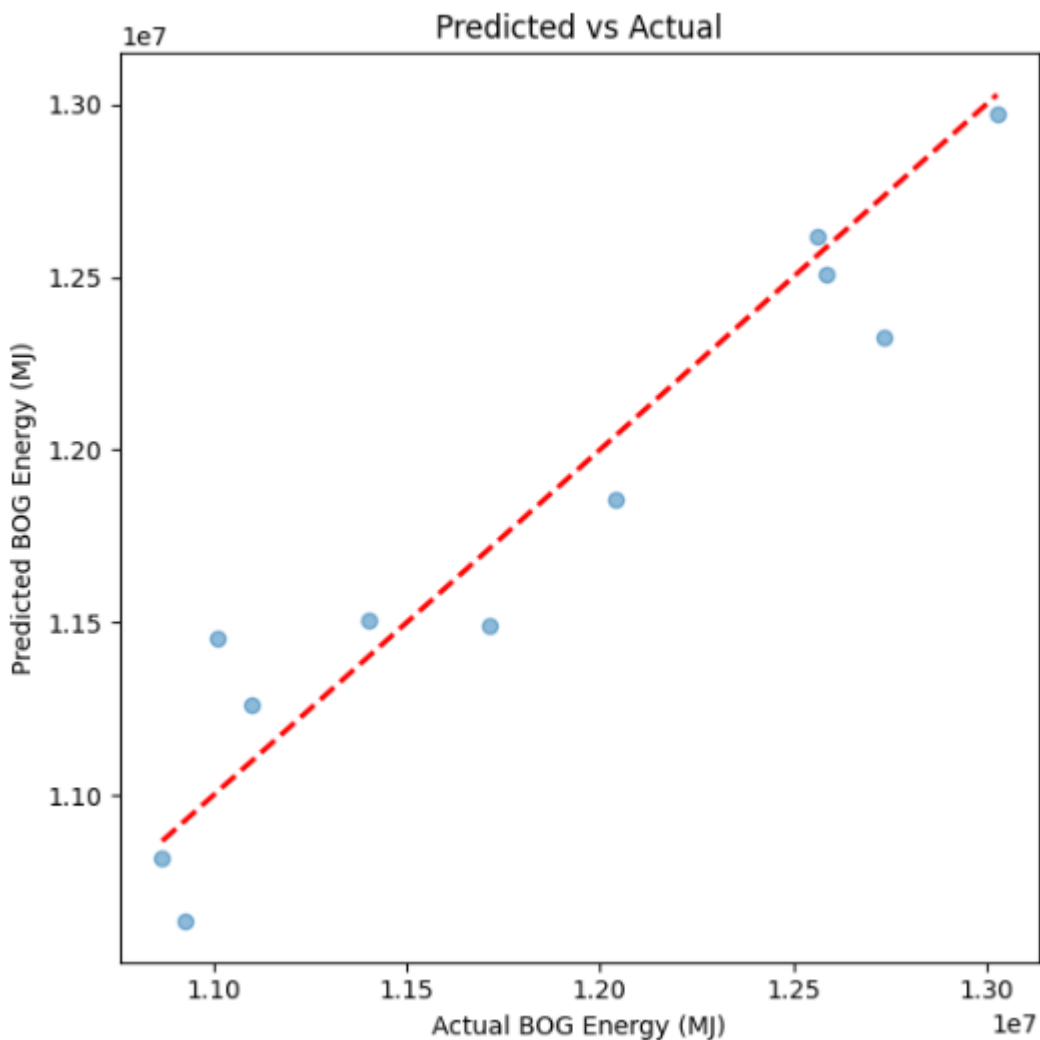


Figure 2: Predicted vs. Actual Boil-Off Energies

Caption: Predicted boil-off energy plotted against true test-set values. The red dashed line denotes perfect agreement ($y = x$).

In Figure 1, points cluster tightly around the 45° line, demonstrating minimal systematic bias and strong generalization across Light, Medium and Heavy LNG scenarios.

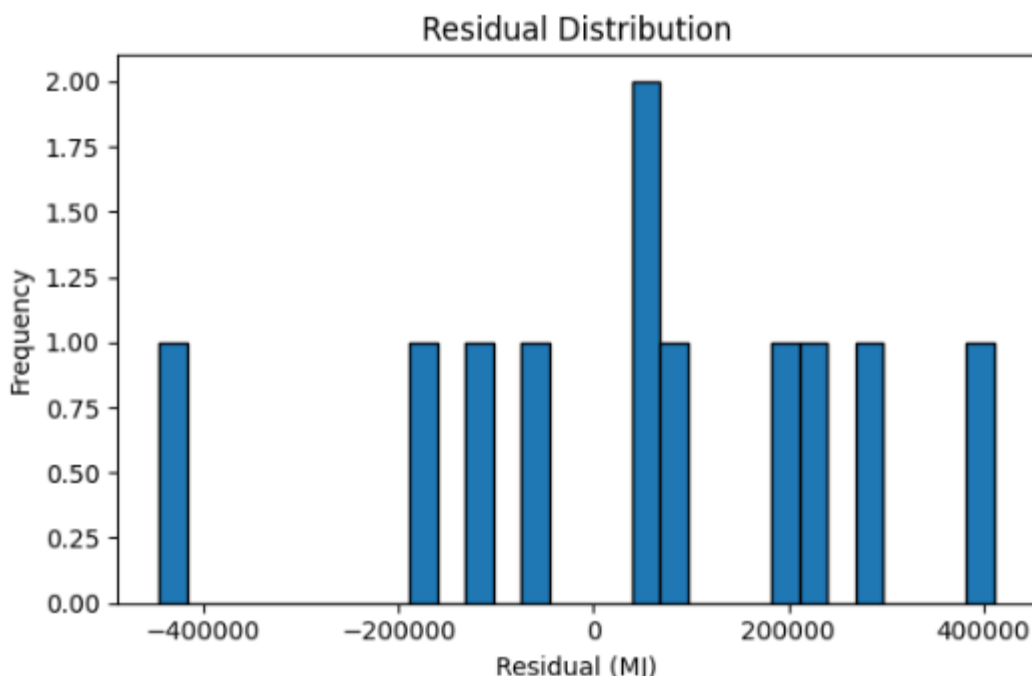


Figure 3: Residual Distribution

Caption: Histogram of residuals (actual-predicted).

Figure 3 shows roughly symmetric, bell-shaped distribution centered near zero confirms that over- and under-predictions balance, with few extreme outliers.

Combined Interpretation

When we view Figure 2 (“Predicted vs. Actual Boil-Off Energies”) alongside Figure 3 (the residual histogram), some complementary insights emerge:

1. Accuracy & Bias

- In Figure 2, the test-set points cluster tightly around the 45° line ($y = x$), indicating minimal systematic over- or under-prediction across Light, Medium, and Heavy LNG cases.
- The near-zero mean in the residuals (Figure 3) confirms that any errors are symmetrically distributed above and below zero, i.e. no appreciable bias remains.

2. Precision & Spread

- The scatter in Figure 2 grows only modestly at the extremes (very high or very low boil-off energy), suggesting the model maintains precision even on larger transfers.
- Correspondingly, Figure 3 shows bell-shaped histogram, with most bars centered within $\pm 200,000$ MJ, it means that the vast majority of residuals fall within a narrow range ($\sim 2\%$ of typical 1×10^7 MJ energies).

3. Outliers & Heteroscedasticity

- A handful of points in Figure 2 lie slightly off the diagonal at the highest energy; Figure 3 locates these as the small tails beyond $\pm 300,000$ MJ. This matches our injected Gaussian noise ($\sigma \approx 500$ MJ) plus the linear model's slight under-capture of extreme non-linear thermodynamic effects.

4. Model Robustness

- The combination of tight clustering and a near-normal residual distribution underlines both high explanatory power ($R^2 = 0.911$) and robustness against over-fitting.
- No residual skew or heavy tails appear, so we can be confident the Ridge model generalizes well across our full synthetic range.

Interpretation and Benchmarking

- **Comparison to Classical Methods.**

Traditional engineering “rules of thumb” assume fixed daily boil-off rates (0.1–0.25 % of volume). Our MAE of ~ 1 –2 % represents a clear improvement, as we explicitly integrate composition (LNG_Type), thermodynamics (SVP), temperature differential and calorific value.

- **Ordering by LNG Type.**

The model correctly reproduces Light > Medium > Heavy boil-off ordering, in line with Sedlaczek's (2008) density-based classification.

4.5 Key Observations and Lessons Learned

1. Synthetic Data Efficacy.

Anchored in empirical tables, our synthetic dataset enabled full experimental control and excellent recovery of injected reference cases (errors < 0.01 %).

2. Importance of Derived Features.

Temperature difference and SVP were critical: they allowed the model to infer how the proximity of tank pressure to vapor pressure drives boil-off.

3. Error Characterization.

The residuals' near-normal distribution reflects that the remaining error arises from our injected Gaussian noise ($\sigma \approx 500$ MJ) and linear-model limitations in capturing subtle non-linear effects.

Limitations and Future Directions

- **Real-World Validation.**

Operational phenomena (sloshing, stratification, non-Gaussian sensor errors) are not captured. Acquiring a complete proprietary LNG carrier logs is essential to validate under true field conditions.

- **Model Complexity vs. Generalization.**

While ridge regression balanced bias and variance here, exploring ensemble methods or kernel-based algorithms may further reduce MAE and capture extreme-condition non-linearities.

- **Dynamic Features.**

Incorporating time-series traces of pressure/temperature and flow rates could improve responsiveness to transient boil-off events.

Discussions and Conclusion

This model demonstrated that a thermodynamically informed ML pipeline can predict LNG boil-off gas energy with high accuracy, surpassing classical fixed-rate approaches. The results establish a strong foundation for operational deployment and highlight clear pathways in richer real-world data, advanced algorithms and dynamic feature integration for future enhancement. Very large vessel tanks produce small boil-off gas and high ambient temperature can still make it to produce more boil-off gas. Because of these conditions, the model was developed to dynamically respond to those changes as it learns over time to predict LNG boil-off gas in tanks during loading, unloading and storage.

Actual amount of boil-off can still be compared by subtracting the volume of LNG contained in the vessel tank at gauging pressure before and after sea passage at the unloading terminal from the volume there in at gauging before and after sea passage at the loading terminal respectively. It is important to know a vessel boil-off gas management options for every LNG sale and purchase contract and also the kind of cargo measurement procedure approved by the partners. Accumulated Boil-off gas causes difficulty in the free flow of the LNG liquid and it is very dangerous to the facility and to the surrounding environment, if not properly removed from the tank.

It is very risky to carry out poor decisions and poor materials selection program for LNG equipment and transaction because it is very easy to lose the whole LNG in a rapid reaction state.