Petroleum 7 (2021) 91–101

Contents lists available at ScienceDirect

Petroleum

journal homepage: www.keaipublishing.com/en/journals/petlm

Predicting saturated vapor pressure of LNG from density and temperature data with a view to improving tank pressure management

David A. Wood

DWA Energy Limited, Lincoln, LN5 9JP, United Kingdom

ARTICLE INFO

Article history:
Received 19 January 2020
Received in revised form
17 March 2020
Accepted 14 April 2020

Keywords:
LNG traded Compositions
LNG SVP relationships
LNG tank pressure influences
Boil-off gas rates
Optimized data-matching prediction

ABSTRACT

Determining the saturated vapor pressure (SVP) of LNG requires detailed thermodynamic calculations based on compositional data. Yet LNG compositions and SVPs evolve constantly for LNG stored in tanks. Moreover, the SVP of the LNG in a tank influences boil-off rates and tank pressure trends. In order to make improved tank pressure control decisions it would be beneficial for LNG tank operators to be made more constantly aware of the SVP of the LNG in a tank. Machine learning models that accurately estimate LNG SVP from density and temperature inputs offer the potential to provide such information. A dataset of five distinct, internationally traded LNG cargoes is compiled with 305 data records representing a range of temperature and density conditions. This can be used graphically to interpolate LNG SVP. However, two machine learning methods are applied to this dataset to automate the SVP predictions. A simple multi-layer perceptron artificial neural network (MLP-ANN) predicts SVP of the dataset with root mean square error (RMSE) = 6.34 kPaA and R^2 = 0.975. The transparent open-box learning network (TOB), a regression-free optimized data matching algorithm predicts SVP of the dataset with RMSE = 0.59 kPaA and R^2 = 0.999. When applied to infill unknown LNG compositions the superior TOB method achieves prediction accuracy of RMSE ~3kPaA and R^2 = 0.996. Predicting LNG SVP to this level of accuracy is beneficial for tank-pressure management decision making.

1. Introduction

The massive expansion and diversification of long-distance trade in the LNG industry in recent years [1] has been associated with steady growth in the LNG shipping fleet, offshore facilities such as floating storage and regasification units (FSRU and FSU) and onshore regasification and storage facilities. Cedigaz [2] forecast international LNG trade to potentially grow by 3.2%/year from 2020 to 2040, and that the share of LNG in net interregional flows is likely to increase from 39% in 2017 to 60% in 2040. However, along the LNG supply chains a certain amount of operational inefficiency exists, particularly in the handling of boil-off gas (BOG) in marine transportation and storage facilities, leading to gas wastage [3,4].

There are a number of industry initiatives aimed at reducing gas cargo waste and emissions associated with LNG shipping. The mandatory Ship Energy Efficiency Management Plan (SEEMP) [5] obliges ship operators to focus on efficiency. The EU Monitoring, Reporting and Verification of CO_2 (EU MRV) made emissions data collection mandatory in January 2018. Additionally, the IMO Data Collection System (IMO DCS) on fuel consumption; data collection became mandatory in January 2019 [6,7]. These relatively recent regulations all focus on fuel and emissions reduction in the shipping sector. The requirements of the SEEMP are gradually motivating the LNG marine industry to adopt more efficient operational practices tailored to modern LNGC facilities, in parallel with developing more efficient engine designs.

In addition to regulatory requirements the commercial terms of LNG charterparty contracts provide incentives for LNGC ship operators to minimize BOG consumption and apply penalties if specified BOG consumption limits are exceeded [8]. LNGC vessel

E-mail address: dw@dwasolutions.com.
Peer review under responsibility of Southwest Petroleum University.

Production and Hosting by Elsevier on behalf of KeAi

https://doi.org/10.1016/j.petlm.2020.04.001

speed, fuel consumption and boil-off limitations all come into play in the charterparty warranties. Nevertheless, operational records of many LNGC and FSRU indicate that unnecessary quantities of boil-off gas are frequently consumed in gas combustion units (GCU) and steam dumps (SD) in actions taken to control LNG tank pressures [3]. Part of the reason for such inefficiency is a general lack of understanding of how tank pressure trends influence boil-off rates and the absolute quantities of BOG production associated with various tank conditions.

It is now clear that LNG tank pressure trends and behaviour are intimately associated with the saturated vapor pressure (SVP) of the LNG in the tank [9]. Indeed, there is a distinct duality of pressure behaviour displayed depending on whether tank pressure is above or below the SVP of the LNG bulk in the tank. Armed with knowledge of the tank pressure versus LNG SVP it is possible to make more informed decisions about when and for how long it is appropriate to run the GCU/SD to mitigate tank pressure increases and stay safely within the maximum allowable relief valve setting (MARVS) specified for the tank. Hence, knowing the SVP of the LNG in a cargo tank at all times is a desirable requirement. The influence of LNG SVP on boil-off rates does not only apply to large-scale LNG tanks. It is also an important factor, related to temperature and pressure, in the boil-off rates incurred in LNG fuel tanks of marine, road vehicles [10,11] and marine bunkering tanks [12]. However, these SVP influences occur on different time scales in large versus small tanks. The large mass of LNG bulk typically present in large LNG storage tanks means that heat exchange processes between LNG and vapor take much longer than in smaller tanks.

Unfortunately, LNG SVP data is not readily available or easily calculated and does not appear as a real-time measured property of LNG tank control systems. Rather, LNG SVP has to be calculated thermodynamically using the interactions of the equations of state that apply to mixtures of light hydrocarbons and nitrogen. In many operating LNG storage and regasification facilities these calculations are usually performed using the Reference Fluid Thermodynamic and Transport Properties (REFPROP) database [13] for specified LNG compositions over a range of temperatures and densities. For example, based on the certified composition of a cargo delivered into a tank as quoted on the accompanying quality notification sheet. However, LNG compositions and their physical properties (pressure and temperature) change progressively during their residency in a tank. This means that SVP of the LNG also evolves, yet tank operators are not immediately informed of the evolving LNG SVP trends on the control system screens. Clearly, a simple method to accurately estimate LNG SVP from LNG temperature and density should provide useful information to tank operators on an ongoing basis. This information would improve their ability to estimate boil-off rates and take more appropriate tank pressure control decisions.

This study proposes and evaluates two supervised machine learning models to predict LNG SVP from LNG temperature data and demonstrates that high prediction accuracy can be reliably achieved. The work is presented in the following order: section 2 describes the composition of internationally traded LNG cargoes and their ageing; section 3 evaluates the density, SVP and temperature distributions of five specific LNG cargo compositions based on a database of 305 data records; section 4 describes two machine learning methods configures to predict SVP from density and temperature data; and, section 5 compares and evaluates the SVP prediction results of the two methods.

2. Commercial LNG compositional range of internationally traded cargoes

LNG cargoes traded internationally display a wide range of compositions. The compositional range is influenced by: 1) the composition of the gas source from which it is derived; 2) the gas specification of the markets into which it is traded; and 3) the aging process that occurs when LNG transportation or storage times are long [4,14]. Some East Asian markets have for many decades preferred LNG of richer compositions (i.e. with methane close $\leq$ 93% and natural gas liquids $\geq$ 6%). On the other hand some LNG producers are only able to easily export lean LNG (i.e., with methane $\geq$ 97.5%). The consuming markets need to adjust the compositions of delivered LNG to match their required gas specifications and adjust for gas interchangeability [15]. This is usually achieved by either adding propane, to make an LNG richer, or adding nitrogen, to make an LNG leaner.

The density of LNG depends on its composition, typically varying between about 405 and 465 kg/m^3, depending on temperature, for internationally traded LNG cargoes. The higher the methane content, the lower the density, and the gradient of density change per unit of temperature is 1.35 kg/m^3/°C [11]. The relationship between SVP and temperature can be mathematically described by the Antoine equation [16].

Table 1 and the supplementary file (see Appendix) provide and compare compositional details of five internationally traded LNGs, shipped on LNGC and stored in onshore LNG storage tanks and floating storage and offloading units. LNG density and SVP are detailed for these five distinct LNG compositions across a temperature range of −150 °C to −162 °C.

While LNG is in transit and/or storage its composition slowly evolves (ages), primarily due to the boil-off quantities of gas removed from the tank to compensate for heat ingress through the tank walls (external) and by running pumps within the tank (internal). This ageing process can be modelled but the compositional changes observed empirically in LNG cargoes show that it is a relatively complex process depending on several factors [17]. As the LNG composition evolves so does its density and SVP at specific temperatures. Tank operators in practice have to deal with a wider range of LNG compositions in their tanks, because although the LNG delivered into their tanks is likely to be of a specific commercial grade, that composition will change the longer that it remains stored and potential becomes mixed with other grades.

Models are proposed that use boil-off rate in tanks to predict changes in LNG composition over time [18,19], and the energy content removed from the liquid phase [20]. However, boil-off rates in tanks vary significantly in most operational conditions, either for onshore or offshore LNG tanks [21]. This makes such model difficult to apply. There is a positive, but non-linear, relationship between boil-off rate in an LNG tank and tank pressure [22]. Tests and empirical observations have revealed that larger tanks tend to experience less boil off in percentage terms than smaller tanks. Additionally, as methane and nitrogen contents of LNG increase the boil-off rate tends to rise. However, experimental studies and models tend not to take into account the empirical observations that show that boil-off rates vary depending upon the magnitude of tank pressure relative to the SVP of the LNG [9].

Some models aim to predict LNG ageing rates in land-based LNG tanks without being driven by boil-off rate [23]. Instead boil-off rate is calculated based on the questionable assumption that a thermodynamic equilibrium is maintained continuously within such tanks [24]. These models then calculate composition changes purely on fluctuating external temperatures that influence the flux of heat ingress into the tanks [25]. Empirical observations, particularly on FSRU and FSU, suggest that to assume a prevailing thermodynamic equilibrium in an LNG tank, with LNG and gas frequently moving in an out of it, is unrealistic. For instance, LNG tank ageing models and rollover prediction models [26] generally assume that LNG and vapor temperatures in a tank are the same or

Table 1
Compositions, densities and SVPs of distinctive LNG cargoes over the temperature range −150 °C to −162 °C calculated thermodynamically with REFPROP [13]. See supplementary file (Appendix) for thermodynamic calculation details of 305 data records providing LNG SVP and density across this temperature interval for each of these LNG compositions. 1 kPa = 0.145038 psi.

Summary Properties Ranges of Five Commercial LNG Cargoes Delivered by LNGC

Summary	LNG Source:	1. Australian LNG (Very Light)		2. Trinidad LNG (Light)		3. Qatar LNG (Ras Laffan)		4. Nigerian LNG (Heavy)		5. Qatar LNG (Ras Laffan Heavy)	
	LNG Composition:	CH4 99.8%; C2H6 0.13%; N2 0.15%		CH4 97.5%; C2H6 2.5%; N2 0.25%		CH4 93.5%; C2H6 5 %; C3H8 1%; N2 0.35 %		CH4 91.5%; C2H6 5.7 %; C3H8 2 %; N2 0.30 %		CH4 89.95%; C2H6 6.33%; C3H8 2.25%; (C4H10)i 0.42%; (C4H10)n 0.65%; C5H12 0.01%; N2 0.39 %	
	LNG Temperature	Liquid Phase Density	LNG Saturation Pressure	Liquid Phase Density	LNG Saturation Pressure	Liquid Phase Density	LNG Saturation Pressure	Liquid Phase Density	LNG Saturation Pressure	Liquid Phase Density	LNG Saturation Pressure
	(oC)	(kg/m^3)	(kPaA)	(kg/m^3)	(kPaA)	(kg/m^3)	(kPaA)	(kg/m^3)	(kPaA)	(kg/m^3)	(kPaA)
Min:	-162.0	405.40	100.75	413.40	95.22	425.54	100.83	437.61	89.58	444.08	100.60
Max:	-150.0	423.57	243.47	431.20	232.58	443.05	237.56	454.68	218.71	461.08	234.58
Mean:	-156.00	414.59	164.34	422.38	156.37	434.38	161.93	446.22	147.10	452.65	160.55

similar at any given point in time. However, empirical observations [9] and simulations [27] reveal that LNG and vapor temperatures differ by as much as 8 °C (vapor hotter than LNG) at given points in time. Temperature differences between LNG and vapor of that magnitude significantly reduce boil-off rates and slow down LNG compositional changes. By how much will depend, to an extent, upon the relationship between tank pressure and the SVP of the LNG. More experimental studies are required that evaluate boil-off rates in relation to tank temperature profiles, pressures and LNG SVP to devise more precise mathematical models for LNG ageing.

3. Density, SVP and temperature relationships for internationally-traded LNG

For the five distinct LNG compositions summarized in Table 1 a total of 305 data records (61 records for each of the five LNG compositions considered) have been compiled with SVP values calculated thermodynamically using temperature and density data (See supplementary file for value details for each data record). The calculated SVP values were generated using REFPROP software [13]. Figs. 1–3 show the relationships between density, temperature and SVP for the LNG compositions considered.

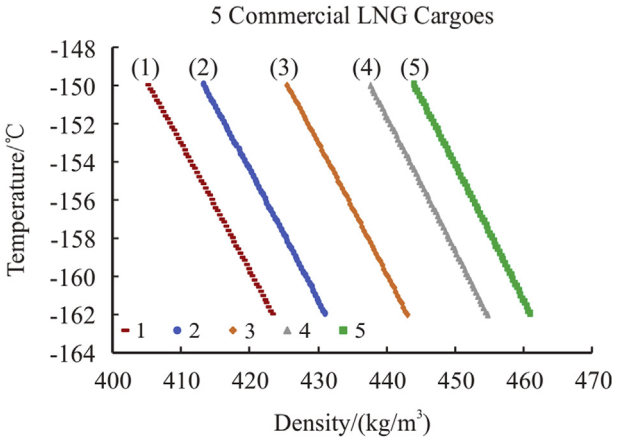


Fig. 1. Temperature versus density for a wide range of commercial LNG compositions. Numbers (1) to (5) identify the five LNG compositions listed in Table 1.

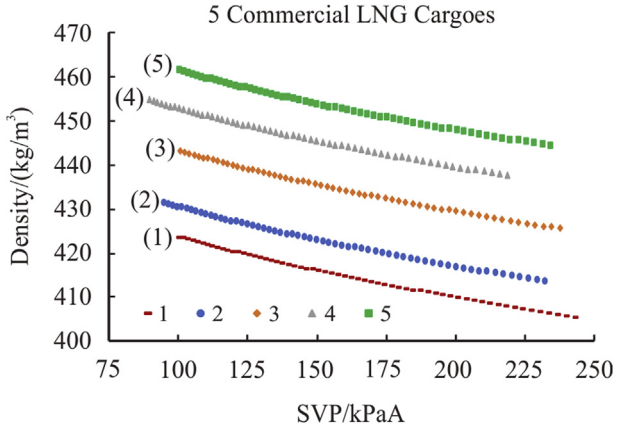


Fig. 2. Density versus SVP for a wide range of commercial LNG compositions. Numbers (1) to (5) identify the five LNG compositions listed in Table 1.

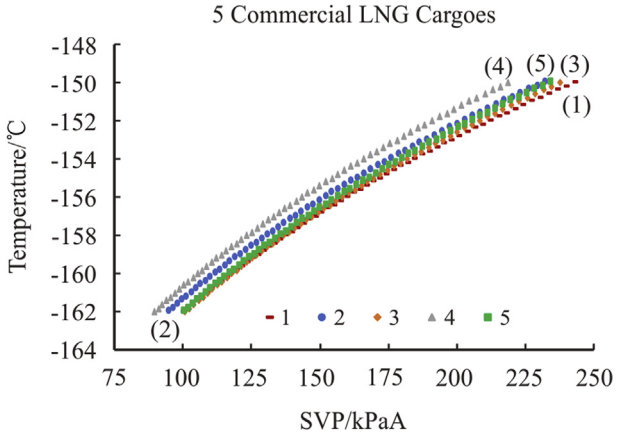


Fig. 3. Temperature versus SVP for a wide range of commercial LNG compositions. Numbers (1) to (5) identify the five LNG compositions listed in Table 1.

During LNG tank operations LNG compositions change over time, on a small scale due to cargo aging [17] and, more significantly, as new cargoes of different composition are loaded into the existing heel during day to day operations of FSRU, FSU and land-

based LNG regasification and storage terminals. Many land-based tanks have level-temperature-density (LTD) tank profiling equipment that provides readily available temperature and density data for all stratified layers of LNG present in a tank at any point in time. However, FSRU, FSU and LNGC typically do not have such equipment fitted and operators have limited information (mainly temperatures at fixed point sensors within their tanks) of the evolving compositions and densities of the LNG in their tanks. Detailed compositional, temperature and density data is available at the time a cargo is loaded and/or offloaded from these floating vessels. Detailed information on density and SVP is typically not available during the day-to-day gas send-out operations and their influences on tank pressure that dominate FSRU and FSU operations. With just LNG temperature data it is possible to estimate LNG SVP crudely to plus or minus 10 to 15 kPaA (Fig. 3). However, that degree of uncertainty in SVP is not of much practical use in LNGC and some FSRU/FSU because the upper safety pressure limit is 250 mbarg (~25kPaA; 1 kPa = 10 mbar)). Knowing the densities of the LNG compositions originally loaded into the LNG tanks, it is possible to estimate LNG density to within about plus or minus 5 kg/m^3.

Figs. 2 and 3 illustrate that temperature and density trends for LNG influence its SVP in different directions. These trends can be easily exploited by machine learning algorithms to routinely predict SVP more accurately than by graphical interpolation given a database of LNG compositional data. Figs. 2 and 3 could be simply used to interpolate graphically LNG SVP for LNG compositions falling within the compiled dataset, including mixtures between the five distinct LNG compositions it considers. However, it is useful for tank control systems to automate such estimates and machine learning methods offer the ability to do this.

4. Method: machine learning networks for SVP predictions

As there are only two independent variables to be considered there are many machine learning algorithms that could be used for the purpose. Here, a basic multi-layer perceptron (MLP) artificial neural network (ANN) to representing non-linear correlation-based machine learning methods. In addition an optimized data-matching method, the transparent open-box (TOB) network [28] is evaluated to represent a machine -learning method that does not rely on any form of correlation between its variables.

4.1. MLP – ANN network model

ANN's are nonlinear machine-learning model that mimics but simplify the neural connections and layers that occur in the human brain [29]. Interconnected artificial neurons (nodes) placed in layers are configured to form a data-processing network. ANN's always contain input and an output layers with one or several hidden layers between them [30]. The number of nodes in the input layer corresponds with the number of input variables (in this study just two: LNG density and temperature). The number of nodes in the output layer corresponds with the number of dependent variables (in this study just one: LNG SVP). The number of hidden layers and nodes they contain can be varied with sensitivity analysis identifying the optimum combination for a specific data set. For this study sensitivity analysis identified that a single hidden layer with five nodes was the optimal configuration.

The values passing from node to node in an ANN are adjusted by weight and bias values which are determined as the network is trained [30]. Additionally, a choice can be made between activation functions to be applied to transform the data values from one network layer for use in the next layer. The selection of the activation function applied in transforming the values between each layer can also be guided by sensitivity analysis [29]. In this study a sigmoidal function applied to the hidden layer and a purely function applied to the output layer generated the best results.

Multilayer perceptron neural networks (MLPNN) [31] are one of the most popular ANN models used. To be effective they require that sufficient data records are allocated to train them [32]. The configuration of the MLPNN developed for this study involves three layers with the one hidden layer possessing five nodes as illustrated in Fig. 4. The value adjustments are made from left (input) to right (output) on a feed-forward basis.

Equation (1) provides the mathematical determination of the value calculated at the jth node in the MLPNN's hidden layer.

$$A_j = B_j + \sum_{j=1}^{m} w_{ij} * X_i \tag{1}$$

Where:

A_j is the weighted summed value of the n input nodes;

X_i is the value at input node i;

w_{ij} is the weight assigned to the connection between the ith input node to the jth (of m) hidden node; and, B_j is the distinct bias term applied to each hidden layer node.

The output value for the j_{th} node in the hidden layer is transformed from the value calculated in equation (1) using equation (2):

$$Y_j = f(A_i) \tag{2}$$

The sigmoidal activation function f applied in equation (2) is expressed by equation (3) and is used to transform the hidden layer output values.

$$f(A) = \frac{1}{1 + e^{-A}} \tag{3}$$

Sensitivity analysis was conducted using other activation functions but the sigmoidal function provided the best results.

The MLP's output layer value $O_k$from the MLPNN is calculated with equation (4)

$$O_k = \left[\sum_{j=1}^{m} w_{jk} * Y_j \right] + B_k \tag{4}$$

Where:

Y_j is the transformed output value for the j_{th} node in the hidden layer;

w_{jk} is the weight between the j_{th} hidden neuron to the one output neuron, k; and, B_k is the bias term associated with the single output neuron. Note that k could be greater than 1 if there are more than one dependent variables.

$O_k$is then adjusted by a simple linear activation function (purelin) to calculate P_k, the predicted SVP value using equation (5).

$$P_k = O_k * 1 \tag{5}$$

All data input to the ANN is transformed into normalized values. Hence, the predicted values need to be denormalized to be expressed in SVP units.

A back-propagation algorithm [33] is applied to optimize the weights and biases as part of the training of the MLPNN model. The objective function optimized by the back-propagation algorithm is the mean squared error (MSE) of the difference between the actual LNG SVP and the predicted LNG SVP for all data records in the training subset (i.e. 266 data records (~87%) spread across the full dataset of 305 data records). The remaining 39 data records (~13%) are allocated to the validation subset for independent evaluation.

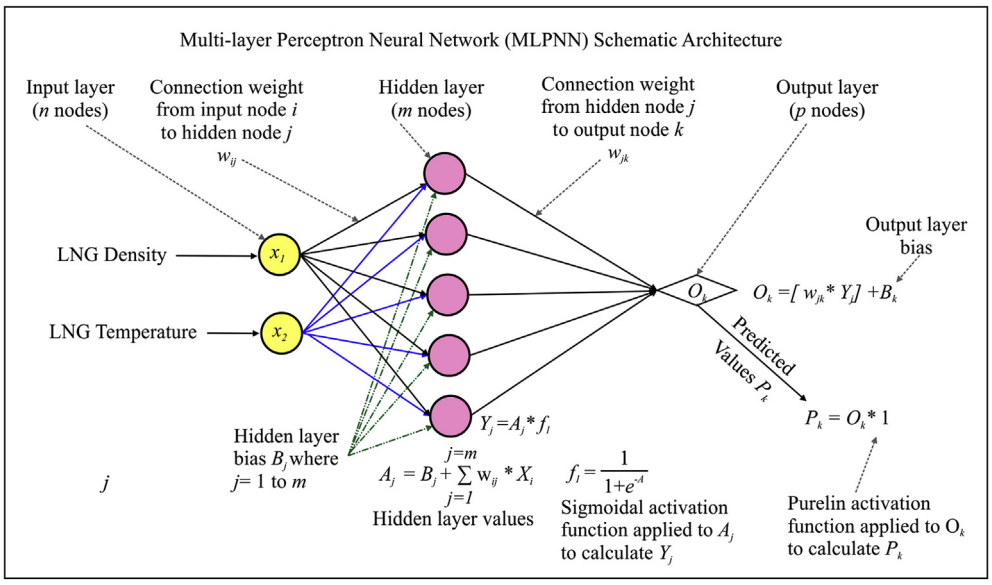


Fig. 4. Schematic illustration of multilayer perceptron neural network (MLPNN) used to predict LNG saturated vapor pressure (SVP) from LNG temperature and density inputs. Just five nodes are involved in a single hidden layer.

The calculation of MSE is achieved using equation (5):

$$MSE = \frac{1}{n}\sum_{i=1}^{i=n}(R_i - P_i) \tag{6}$$

Where:

R_i is the real thermodynamically determined LNG SVP value for a given training subset data record i;

P_i is the calculated/predicted LNG SVP value for training subset data record i;

and, n is the total number of training subset data records.

Once the MLPNN is trained and validated it is applied to predict LNG SVP for all 305 data records in the dataset. It is also then available to predict LNG SVP on an unsupervised basis for other unknown LNG samples based on their temperatures and densities.

4.2. TOB – Optimized data matching model

The TOB optimized data matching method [28] differs fundamentally from ANN in that it involves no correlations, regressions or hidden layers. It transparently provides effective supervised machine learning predictions for large and small datasets [34]. The TOB method is described in detail in the studies cited and is not repeated here. Fig. 5 provides a flow diagram summarizing the TOB calculation sequence adapted for predicting LNG SVP from density and temperature data. The dataset is divided into training, tuning and testing subsets (Fig. 5 steps 1 to 6). Each data record in the tuning subset is matched against all the records in the training subset. The allocation of data records to the subsets (training, tuning and testing) is conducted in a semi-random fashion without replacement. It is not fully random because that could lead to clustering across the value distribution in some subsets. The subsets are therefore selected randomly from sections spread across the full value range of the SVP distributions to ensure that all subsets include a representative spread of data records. Executing the algorithm using several different subset samples selected in this way leads to repeatable results.

The squared differences (or errors) of the independent variables (*VSE*) for each of *J* tuning-subset data records versus the *K* training-subset data records are calculated in TOB Step 7 (Fig. 5) using equation (7):

$$VSE(X)_{jk} = \left[X_k(tr) - X_j(tu)\right]^2 \tag{7}$$

Where:

$X_k(tr)$ = variable X *value* for the kth training-subset record
$X_j(tu)$ = variable X for the jth tuning-subset record
$VSE(X)_{jk}$ = variable-squared error (VSE) for variable X of the jth tuning-subset record versus variable X of the kth training-subset record.

$VSE(Xn)_{jk}$ = the variable-squared error (*VSE*) for variable *Xn* for the jth tuning-subset record versus the kth training-subset record. $\sum VSE_{jk}$ computes the weighted sum of the computed *VSE* values combining the squared errors of all the independent variables applying equation (8)

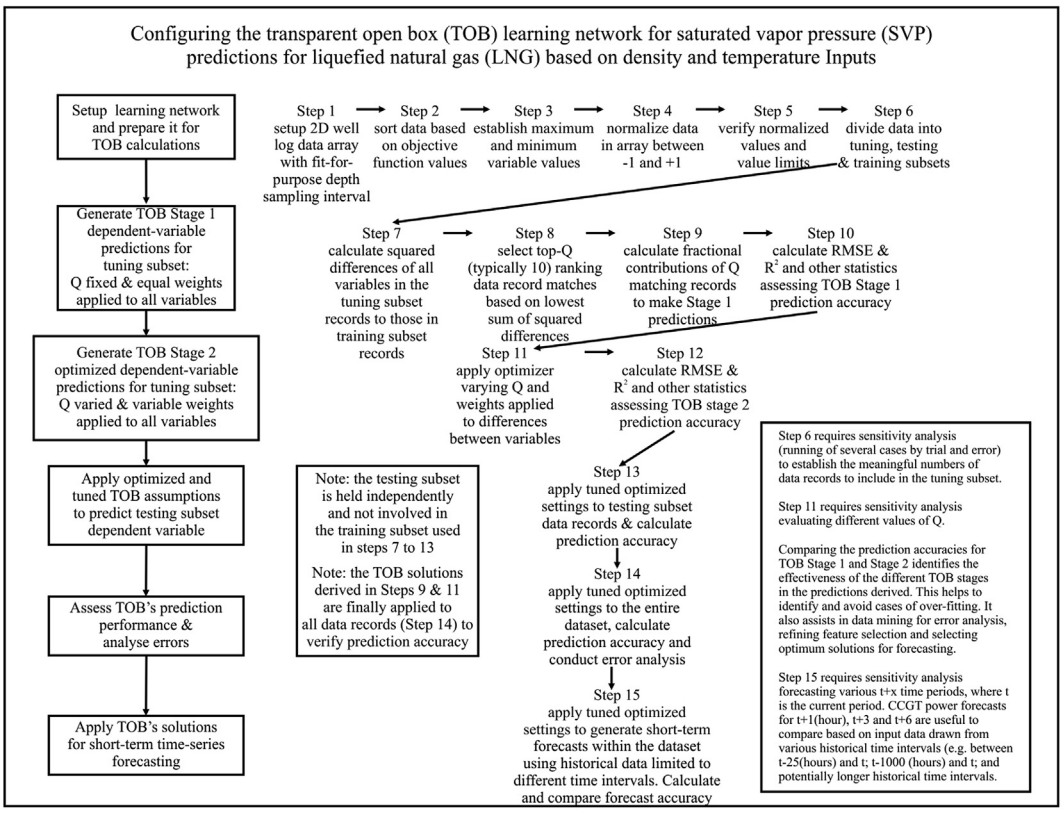
$$\sum VSE_{jk} = \sum_{n=1}^{n=N+1} VSE(Xn)_{jk} * (Wn) \tag{8}$$

Where:

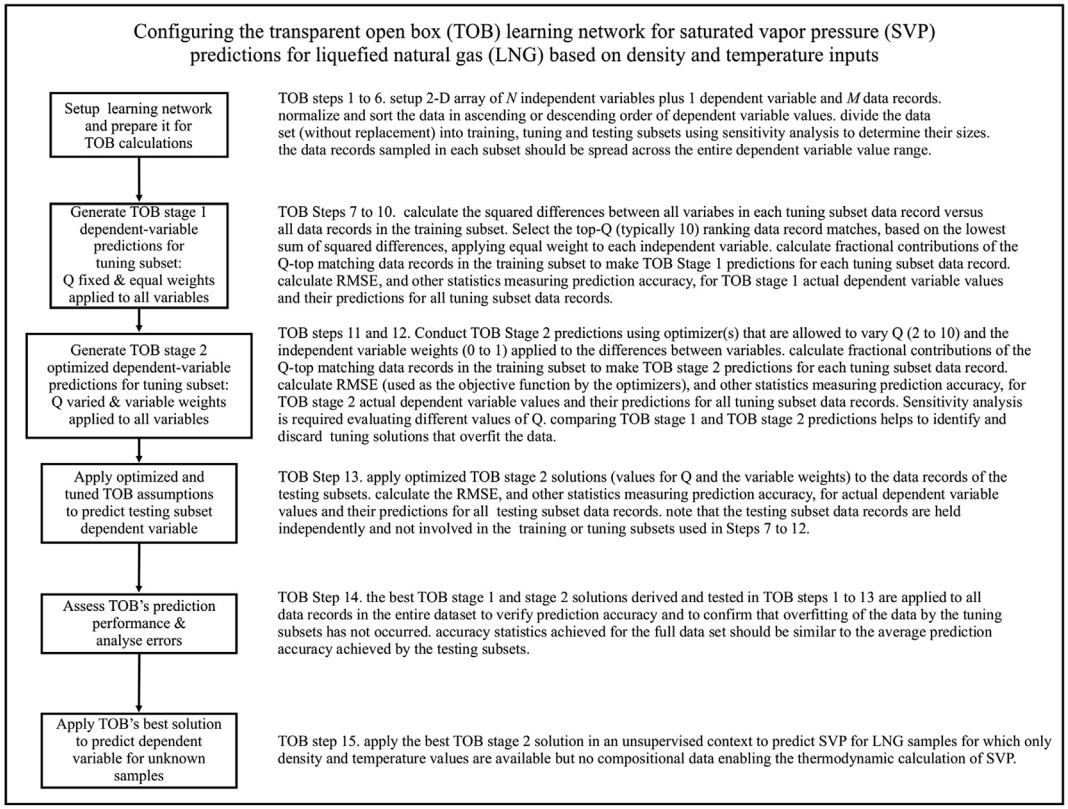
Wn = weights ($0 < Wn \leq 1$) applied to the calculated *VSE* for all variables involved in the prediction (i.e., $N+1$). Each weight is set to a constant value (e.g. 0.5 or 1.0) in TOB Stage 1. This means that TOB Stage 1 establishes the high-ranking data record matches initially without any biases attached to any of the independent variables.

The best matching data records are those with the lowest calculated $\sum VSE$ according to equation (7). $\sum VSE$ values are used to identify the top-ten matching records ($Q = 10$) in the training subset for each record in the tuning subset (Step 8 Fig. 5).

The contribution fraction to the dependent variable prediction from each of the top-Q matching data records (where Q = 10 for TOB Stage 1) is determined using equations (9)–(11).



(a)



(b)

Fig. 5. Transparent open box algorithm workflow with steps configured for predicting SVP of LNG from density and temperature.

$$f_{jq} = \sum VSE_{jq} \Big/ \left[\sum_{r=1}^{r=Q} \sum VSE_{jr} \right] \tag{9}$$

Where:

q = the qth of Q top-ranking training-subset records from the training subset for the jth tuning subset record.
r = the rth of Q top-ranking training-subset records from the training subset for the jth tuning subset record.
f_q = the contribution fraction calculated for the qth of Q top-ranking records for the jth tuning subset record. Equation (10) expresses an essential constraint applied to these contribution values:

$$\sum_{q=1}^{q=Q} f_q = 1 \tag{10}$$

This normalizes the f_q values to sum to 1. Q always equals 10 for TOB stage 1.

The data record in the training subset with the highest-ranking match is the one with the lowest $\sum VSE_{jk}$and f_q values. This means that in order to assign the greatest contribution to the prediction of the dependent variable to the best matching training subset record with the jth tuning-subset record $(1 - f_q)$ multipliers need to be applied using equation (11))

$$(X_{N+1})_j^{predicted} = \sum_{q=1}^{q=Q} \left[(X_{N+1})_q * \left(1 - f_q\right) \right] \tag{11}$$

Where:

$(X_{N+1})_q$ = dependent variable for the qth training-subset record (i.e., one of Q best-matching records).

$(X_{N+1})_j^{predicted}$ = TOB-stage-one predicted-dependent-variable value for the jth tuning-subset record.

These calculations provide the TOB Stage 1 predictions (Step 10 Fig. 5).

The same calculation equations (7)–(11) are used for TOB Stage 2 prediction but Q and Wn are allowed to vary (unlike in TOB Stage 1 where they are held constant): (0< Wn ≤ 1) and (2 ≤ Q<=10). TOB stage-2 $\sum VSE_{jq}$values are recomputed with equation (9) applying the available range of Q values (2< Q ≤ 10) in each iteration of the optimizer. Optimizers are applied to find the lowest root mean square error (RMSE) between the actual and predicted dependent variable values (Steps 11 and 12 Fig. 5).

For small data sets (up to a few thousand data records) TOB Stage 2 optimization can be efficiently conducted using Excel's Solver optimizers [35]. However, a memetic firefly optimizer [34] provides the flexibility to run the optimization in a fully coded version independently of Excel. Indeed, other non-linear optimizers of the gradient-reduction type and evolutionary algorithms with good local and global search capabilities can be easily adapted for use with the TOB algorithm. An independent testing subset is used to verify that the optimum solutions derived by applying the two stages of the TOB algorithm to the tuning subset provide dependent-variable predictions that are statistically valid when applied to the dataset more generally. Finally, the best TOB Stage 2 solution is applied to predict the dependent variable values in all the dataset records.

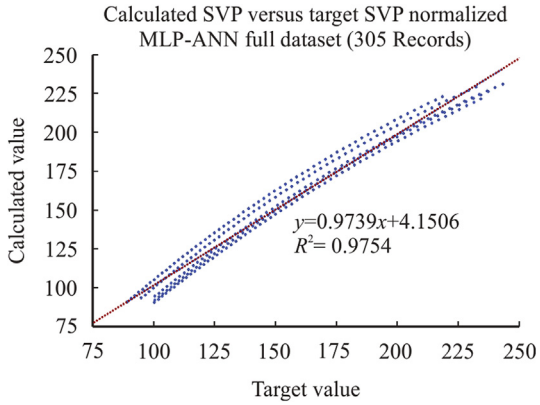


Fig. 6. LNG SVP prediction accuracy for the MLP-ANN model. The red dotted line is the linear regression line through the data point and the formula for that line is shown in the lower right corner of the graph.

5. Results

5.1. Prediction results of the MLP-ANN model

The trained MLP-ANN model applied to all 305 data records in the LNG SVP data set predicts SVP with an accuracy of RMSE = 6.34 kPaA and R^2 = 0.975 (Fig. 6).

The weights and biases of the trained MLP-ANN model are listed in Table 2.

It is apparent from Fig. 6 that the ANN model provides a better fit between actual and predicted LNG SVP in the middle range of values for some compositions with poorer fits at the lower and upper end of the SVP range covered by the dataset, and vice versa for other LNG compositions.

The weights and biases listed in Table 2 can be entered into equation (12) along with the normalized data values for LNG density and temperature to calculate the normalized LNG SVP for each data record.

$$P_k = \left[\sum_{j=1}^{m} \left(\frac{1}{1 + e^{-\left(\left(X_1 \cdot w_{1j} + B_j \right) + \left(X_2 \cdot w_{2j} + B_j \right) \right)}} \right) \cdot W_{jk} \right] + B_k \tag{12}$$

Where the symbols in equation (12) are those used in Fig. 4 and defined in terms of equations (1)–(4), and X_1 is the normalized input value for LNG density and X_2 is the normalized input value for LNG temperature.

Equation (12) takes this form with relatively few terms, because there are only two input nodes (n = 2), five nodes in the hidden layer (m = 5), the input to hidden layer activation function is sigmoidal (equation (3)), the output activation function is purelin and there is just one dependent variable (k = 1). The equation is relatively easy to code in an Excel sheet or other software to calculate SVP from density and temperature inputs along with a normalization and denormalization adjustment.

5.2. Prediction results of the TOB model

The TOB learning network achieves much more accurate predictions for LNG SVP from this dataset than the MLP-ANN model. Indeed, the TOB Stage 1 predictions using equal error weights (Wn = 0.5) and Q = 10 achieve prediction accuracy of RMSE = 0.8942 kPaA and R^2 = 0.9996 (Fig. 7). TOB Stage 2 predictions further improve on that accuracy with the best-case optimized solution (Q = 2; $W_{temperature}$ = 1; $W_{density}$ = 8.62 E−05)

Table 2
Weights and biases of trained MLP-ANN applied with equation (12).

Weights and Biases for Trained MLP-ANN Model to Predict LNG SVP						
J=5	Node 1	Node 2	Node 3	Node 4	Node 5	
Input to Hidden layer						
Bias (Bj)	-0.687307	-0.418680	-0.448536	-0.046981	-0.378652	
Weight (W1J)	-0.567330	-0.067332	0.213771	-0.245953	-0.059916	(Density)
Weight (W2J)	1.268382	-0.499609	1.083432	-0.664911	-0.112893	(Temperature)
Hidden to Output layer						**Bias (Bk)**
Weights	1.360599	-0.554968	1.003513	-0.736214	-0.253871	-0.116032854133084

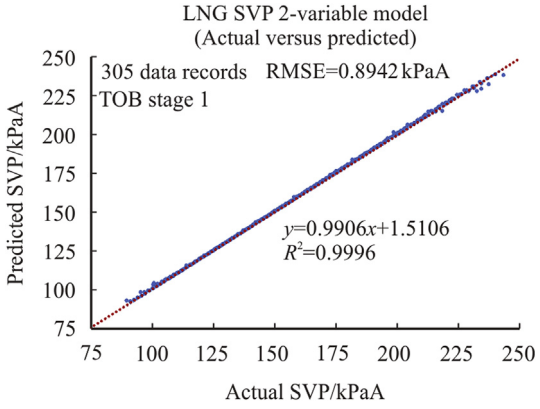


Fig. 7. LNG SVP prediction accuracy for the TOB Stage 1 model. The red dotted line is the linear regression line through the data point and the formula for that line is shown in the lower right corner of the graph.

achieving prediction accuracy of RMSE = 0.5882 kPaA and R^2 = 0.9998 (Fig. 8). The very high accuracy of the TOB Stage 1 model is improved by reducing the best matching records used to just the best two ($Q = 2$) and by placing much more weight on the temperature match than the density match. Although the weight applied to density errors is very small, it is not zero, and that means it does play a crucial role in the prediction accuracy achieved. For instance, in matching data records with identical LNG temperatures that small density error weight would mean that the matching record with the closest density match would be ranked higher.

It is important to recognise that the TOB algorithm involves a two-stage process. The data matching and TOB Stage 1 analysis are conducted with equal weights applied to all input variables. This means that if an input variable is assigned a zero weight in the TOB optimization stage 2 that does not mean that variable has had no influence at all in the prediction, because it has been involved in the initial selection of closely matched data records. Moreover, a very low (non-zero) weight applied to a variable in the TOB stage 2 is often very useful to the optimization process. This is so in cases where closely matching data records have exactly the same values for some variables. A small weight applied to a variable with distinct values can then be employed by the optimizer to distinguish between them.

6. Testing and applying the machine learning models

A number of interpolated mixtures between adjacent LNG compositions displayed in Table 1 and Fig. 1 are calculated to establish a set of "unknown SVP" data points with infill mixed LNG temperatures and densities located between the values of the 5 known LNGs (triangles in Fig. 9). These mixtures are 0.50:0.5, 0.75:0.25 and 0.25:0.75 vol mixes of the adjacent known LNGs (circles). The assumption is made that temperature and density of the mixtures will vary on a pro rata basis between the end point compositions mixed. This is clearly an approximation, but the parallel linear trends of the five known LNGs (Figs. 1 and 9) suggest

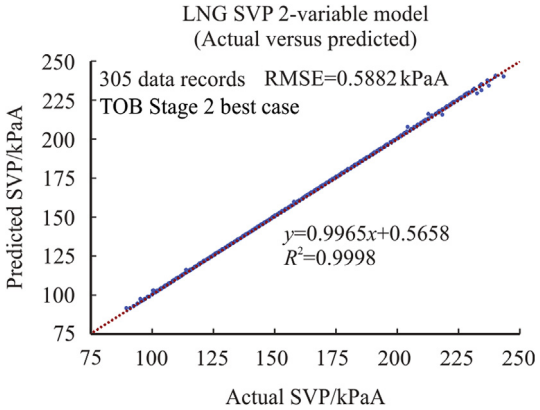


Fig. 8. LNG SVP prediction accuracy for the TOB Stage 2 model. The red dotted line is the linear regression line through the data point and the formula for that line is shown in the lower right corner of the graph.

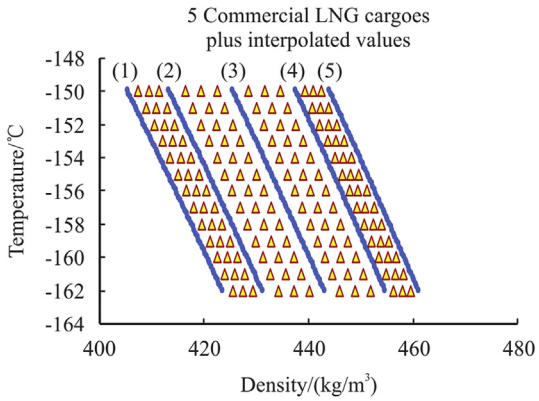


Fig. 9. Interpolated temperature and density properties (orange triangles) for mixtures of the five distinct LNG cargoes (blue circles) for which there is thermodynamically calculated SVP values.

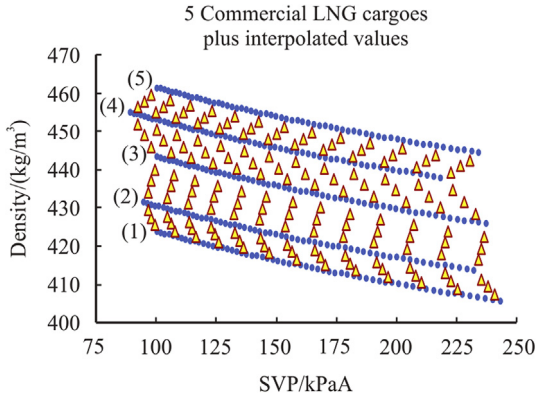


Fig. 10. Interpolated density and SVP properties (orange triangles) for mixtures of the.

that this is not an unreasonable assumption.

A further assumption is made in Fig. 10 that SVP and density of the mixtures will vary on a pro rata basis between the end point compositions mixed. This is also clearly an approximation, but the parallel near-linear trends of the five known LNGs (Figs. 2 and 10) suggest that this is not an unreasonable assumption. However, the temperature versus SVP trends (Fig. 3) is non-linear and the five known LNG compositions are not distributed as systematically in that figure as for the density versus SVP distributions. This suggests that interpolated SVP values from Fig. 10 can only be considered as approximate for the mixtures created. However, this set of approximate unknown LNG mixtures represents a useful dataset with which to test the prediction capabilities of the trained MLP-ANN and TOB models for temperature-density data for which SVP is not known.

Five distinct LNG cargoes (blue circles) for which there is thermodynamically calculated SVP values. The interpolated SVP values for this mixed LNG compositions should only be considered as approximations of the thermodynamic SVP values for those mixtures.

There are 156 infill ("unknown SVP" data points (triangles) in Figs. 9 and 10. In relation to how close the estimated SVPs are to the approximate interpolated SVP values (Fig. 10), the MLP-ANN model trained on the known values (circles, Fig. 10) achieves accuracy of RMSE = 4.306 kPaA and R^2 = 0.9899 (Fig. 11).

Notice that as with the dataset known SVP records the MLP-ANN SVP predictions are better in the middle of the SVP value range than at the lower or upper ends of that range. There is a risk in such conditions that the trained MLP-ANN model may overfit the infill unknown data.

In relation to how close the estimated SVPs are to the approximate interpolated SVP values (Fig. 10), the TOB Stage 1 model trained on the known values (circles, Fig. 10) achieves accuracy of RMSE = 2.969 kPaA and R^2 = 0.9957 (Fig. 12). The TOB Stage 2 model does not improve on that, achieving accuracy of RMSE = 3.376 kPaA and R^2 = 0.9939 for the infill test samples. The distance between the known SVP values of the five LNGs (Figs. 9 and 10) places some accuracy constraints on samples between those known data records. Each of the infill samples (triangles in Figs. 9 and 10) can only be matched with known data records (circles in Figs. 9 and 10) to achieve predictions using the trained TOB model. The TOB model does not overfit the data.

Both MLP-ANN and TOB models produce credible SVP predictions for samples spread across the LNG density/temperature feasible space. Overall, the TOB model achieves the best accuracy (RMSE ~ 3 kPaA) and is reliable across the entire LNG SVP range evaluated and not prone to overfitting. The MLP-ANN model, on the other hand, predicts with an accuracy of RMSE >4 kPaA and is less reliable at the lower and upper end of the SVP range. The results and analysis presented suggests that the TOB algorithm offers a more accurate method than the MLP-ANN algorithm to estimate the SVP of LNG more accurately from temperature and density input data. The prediction performance accuracy of the TOB method is very high with respect to calculated SVP for a wide range of LNG compositions derived from Reference Fluid Thermodynamic and Transport Properties (REFPROP) system that calculates SVP of LNG thermodynamically using phase information and equations of state.

To the authors knowledge there are no previously published machine learning methods that attempt to predict LNG SVP from temperature and density data only. The main practical benefit of access to accurate estimates of LNG SVP are for the operators of LNG tanks in storage and regasification facilities. However, being able to predict SVP with reasonable accuracy from temperature and density data also has other academic and industrial applications in the

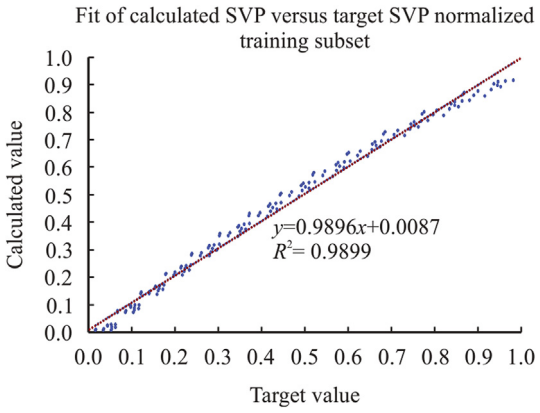


Fig. 11. Normalized predicted versus actual LNG SVP for the unknown infill data records applying the MLP-ANN model trained only with known data records. The red dotted line is the linear regression line through the data point and the formula for that line is shown in the lower right corner of the graph.

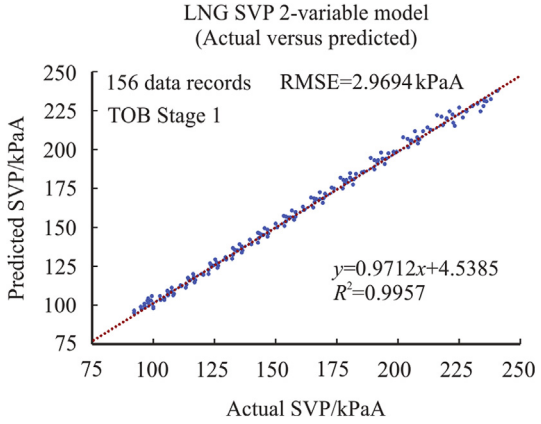


Fig. 12. Predicted versus actual LNG SVP for the unknown infill data records applying the TOB Stage 1 model trained only with known data records. The red dotted line is the linear regression line through the data point and the formula for that line is shown in the lower right corner of the graph.

context of various LNG tank simulation studies and LNG weathering models.

To calculate SVP of LNG based on equations of state, a substantial amount of input information is required for the LNG under consideration, with respect to the phases involved and their chemistry. In many LNG operating contexts such information is not readily available, just temperature and density. The advantage of the proposed method is that it can be used with just LNG temperature and density information to provide meaningfully accurate estimates of SVP, based on an underlying dataset of thermodynamically derived LNG SVP measurements. It is not proposed that these prediction algorithms should act as replacements for thermodynamic calculations via equations of state, rather that they offer practical alternatives in circumstances where limited or no phase/compositional information is readily available for the LNG being handled.

One way to further improve the accuracy of the TOB model is to add thermodynamically calculated SVP values for mixed LNG cargoes residing between the five LNG cargoes used in this study. With more infill known samples the data matching would be improved as the space between the circles in Figs. 9 and 10 would be reduced. An advantage of the TOB model is that as more known data becomes available it can be added into the TOB training subset making more accurate matching possible.

Both of the machine learning methods could be applied in real time in operational situations to provide LNG facility tank operators with constant insight to the SVP of the LNG existing, as a single homogenized layer or in stratified layers, within a tank. Such information would provide them with the necessary insight to make improved tank pressure management decisions.

7. Conclusions

The saturated vapor pressure (SVP) of liquefied natural gas (LNG) is a key attribute to be aware of for LNG stored within a tank in order to better control and respond to tank pressure changes. Whereas, detailed thermodynamic calculations based upon LNG compositions are required to provide direct SVP information, machine learning methods can reliably estimate SVP from density and temperature data for a dataset that encompasses the range of LNG compositions that are typically traded in international markets. Comparison of an artificial neural network model (MLP-ANN, correlation based) and an optimized data matching network model (TOB, correlation-free) indicate that the TOB model provides superior SVP predictions than the MLP-ANN model for the compiled LNG dataset. It achieves accuracy of root mean square error (RMSE) = 0.59 kPaA for predicting SVP on a supervised bases for compositions in the compiled dataset. On the other hand, for infill unknown compositions it achieves SVP prediction accuracy of RMSE ~3kPaA. Such prediction accuracy is considered reliable and beneficial for tank pressure management purposes, the main objective of this study, but should also be useful for some LNG tank simulation studies and aging models. Such machine-learning methods are designed to complement, not replace or outperform, thermodynamic calculations of LNG SVP based on compositional data and equations of state. Prediction accuracy of the TOB model could be further improved by adding thermodynamically calculated SVP for additional LNG compositions to the dataset compiled here.

Declaration of competing interests

The authors declare that they have no conflict of interests.

Appendix

A supplementary Excel file is available with 305 data records listing LNG SVP and density across the temperature interval −150 °C to −162 °C for the five internationally traded LNG compositions evaluated in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.petlm.2020.04.001.

References

[1] GIIGNL, The LNG industry annual report 2019. International Group of Liquefied Natural Gas Importers (GIIGNL), 2019, 56 pages, https://giignl.org/sites/default/files/PUBLIC_AREA/Publications/giignl_annual_report_2019-compressed.pdf. (Accessed 15 January 2020). Accessed.

[2] CEDIGAZ, World LNG Outlook 2019, CEDIGAZ the International Association for Natural Gas, 2019, p. 20. https://www.cedigaz.org/category/analysis/. (Accessed 15 January 2020). Accessed.

[3] M. Kulitsa, D.A. Wood, Part 1: improved monitoring onboard FSRUs is required to enhance operating performance and cut cargo loss, LNG Journal (October Issue) (2017) 22–24.

[4] M. Kulitsa, D.A. Wood, Enhanced application for FSRU recondensing equipment during periods of low or No gas send out to minimize LNG cargo losses, Petroleum 4 (4) (2018) 365–374, https://doi.org/10.1016/j.petlm.2018.01.002.

[5] MEPC, Guidelines for the Development of a Ship Energy Efficiency Management Plan (SEEMP). Resolution MEPC.282(70), Adopted by Marine Environment Protection Committee vol. 28 October (2016).

[6] DNV-GL, EU MRV, IMO DCS. https://www.dnvgl.com/maritime/insights/topics/EU-MRV-and-IMO-DCS/index.html, 2019. (Accessed 15 January 2019). Accessed.

[7] ICS-Shipping, European union monitoring, reporting and verification (EU MRV) regulation. International chamber of shipping guidance. https://www.ics-shipping.org/docs/default-source/resources/ics-guidance-on-eu-mrv.pdf?sfvrsn=10, 2015. (Accessed 16 December 2019) accessed, 13.

[8] S. Mackin, R. Howells, In-depth analysis - LNG: speed, consumption and boil-off warranties. Clyde & Co insight, 7 pages, https://www.clydeco.com/insight/article/in-depth-analysis-lng-speed-consumption-and-boil-off-warranties, 2020. (Accessed 15 January 2020). Accessed.

[9] M. Kulitsa, D.A. Wood, Soft metal blanket with optional anti-sloshing conceptual designs to improve pressure control for floating and land-based liquefied natural gas tanks, Advances in Geo-Energy Research 3 (4) (2019) 424–447, https://doi.org/10.26804/ager.2019.04.09.

[10] Q.-S. Chen, J. Wegrzyn, V. Prasad, Analysis of temperature and pressure changes in liquefied natural gas (LNG) cryogenic tanks, Cryogenics 44 (2004) 701–709.

[11] K. Li, G. Wang, C. Liu, Study on calculation of liquid level and storage of tanks for LNG-fuelled vessels, IOP Conf. Ser. Earth Environ. Sci. 111 (2018), 012030, https://doi.org/10.1088/1755-1315/111/1/012030, 8 pages.

[12] Y. Shao, Y. Lee, H. Kang, Dynamic optimization of boil-off gas generation for different time limits in liquid natural gas bunkering, Energies 12 (2019) 1130, https://doi.org/10.3390/en12061130.

[13] NIST, Reference Fluid thermodynamic and Transport properties database (REFPROP): version 10. https://www.nist.gov/programs-projects/reference-fluid-thermodynamic-and-transport-properties-database-refprop, 2020. (Accessed 8 January 2020) accessed.

[14] J.M. Shah, J.J. Aarts, Effect of weathering of LNG in storage tanks, in: K.D. Timmerhaus (Ed.), Advances in Cryogenic Engineering, Springer Science & Business, 1995, pp. 253–260.

[15] D.A. Wood, S. Mokhatab, Natural gas interchangeability in focus as sources of LNG widen, LNG Journal (February) (2007) 14–18.

[16] F. Yang, Measurement method of liquid level in LNG tank[J], Tianjin chemical 4 (4) (2014) 54–55, 27.

[17] D.A. Wood, M. Kulitsa, Weathering/ageing of LNG cargoes during marine Transport and processing on FSU and FSRU, J. Energy Resour. Technol. 140 (10) (2018), 102901, https://doi.org/10.1115/1.4039981 (11 pages).

[18] A. Benito, Accurate Determination of LNG Quality Unloaded in Receiving Terminals: an Innovative Approach, GERG Academic Network Event, Brussels, Belgium, 2009, pp. 1–23.

[19] M. Miana, R. Del Hoyo, V. Rodrigálvarez, J.R. Valdés, R. Llorens, Calculation models of Liquefied Natural Gas (LNG) weathering during ship transportation, Appl. Energy 87 (2010) 1687–1700.

[20] E. Adom, S.Z. Islam, X. Ji, Modelling of boil-off gas in LNG tanks: a case study, Int. J. Eng. Technol. 2 (4) (2010) 292–296.

[21] D. Dobrota, B. Lalić, I. Komar, Problem of boil - off in LNG supply chain, Trans. Marit. Sci. (2013) 91–100, 02.

[22] G.G. Dimopoulis, C.A. Frangopoulos, A dynamic model for liquefied natural gas evaporation during marine transportation, Int. J. Therm. 11 (3) (2008)

123–131.
[23] L.A. Pellegrini, S. Moioli, F. Brignoli, C. Bellini, L.N.G. Technology, The weathering in above-ground storage tanks, Ind. Eng. Chem. Res. 53 (10) (2014) 3931–3937.
[24] C. Migliore, C. Tubilleja, V. Vesovic, Weathering prediction model for stored liquefied natural gas (LNG), J. Nat. Gas Sci. Eng. 26 (2015) 570–580.
[25] C.M. Capello, Modelling the Weathering Process of Stored Liquefied Natural Gas (LNG). PhD Thesis, Imperial College London, 2016, p. 305.
[26] A. Hubert, S. Dembele, P. Denissenko, J. Wen, Predicting liquefied natural gas (LNG) rollovers using computational Fluid dynamics, J. Loss Prev. Process. Ind. 62 (2019), https://doi.org/10.1016/j.jlp.2019.103922.
[27] C. Migliore, A. Salehi, V. Vesovic, A non-equilibrium approach to modelling the weathering of stored Liquefied Natural Gas (LNG), Energy 124 (2017) 684–692.
[28] D.A. Wood, transparent, Open-Box learning network provides insight to complex systems and a performance benchmark for more-opaque machine learning algorithms, Advances in Geo-Energy Research 2 (2) (2018) 148–162, https://doi.org/10.26804/ager.2018.02.04 (2018).
[29] F. Fausett, Fundamentals of Neural Networks: Architectures, Algorithms and Applications, Prentice-Hall, Englewood Cliffs, NJ, 1994.
[30] S. Haykin, Neural Networks: A Comprehensive Foundation, first ed., 1994 (Saddle River, NJ, USA).
[31] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Network. 2 (1989) 359–366, https://doi.org/10.1016/0893-6080 (89) 90020-8.
[32] K. Hornik, Approximation capabilities of multilayer feedforward networks, Neural Network. 4 (1991) 251–257, https://doi.org/10.1016/0893-6080(91)90009-T.
[33] R. Battiti, First and second order methods for learning: between steepest descent and Newton's method, Neural Comput. 4 (2) (1992) 141–166.
[34] D.A. Wood, Transparent open-box learning network provides auditable predictions for coal gross calorific value, Modeling Earth Systems and Environment 5 (2) (2019) 395–419, https://doi.org/10.1007/s40808-018-0543-9.
[35] Frontline Solvers, Standard Excel Solver - Limitations of Nonlinear Optimization, 2020 (accessed: 13 January 2020), https://www.solver.com/standard-excel-solver-limitations-nonlinear-optimization.