Journal of CO2 Utilization 95 (2025) 103072

Contents lists available at ScienceDirect

Journal of CO2 Utilization

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

Machine learning integration in thermodynamics: Predicting CO_2 mixture saturation properties for sustainable refrigeration applications

Carlos G. Albà [a,b], Ismail I.I. Alkhatib [b], Lourdes F. Vega [b,*], Fèlix Llovell [a,*]

[a] Department of Chemical Engineering, ETSEQ, Universitat Rovira i Virgili (URV), Campus Sescelades, Av. Països Catalans 26, Tarragona 43007, Spain
[b] Research and Innovation Center on CO2 and Hydrogen (RICH Center) and Department of Chemical and Petroleum Engineering, Khalifa University of Science and Technology, PO Box 127788, Abu Dhabi, United Arab Emirates

ARTICLE INFO

Keywords:
Global Warming
CO_2-based Refrigerants
Machine Learning
Polar soft-SAFT
COSMO-RS

ABSTRACT

The need for sustainable alternatives in refrigeration has grown as Europe enforces mandates on avoiding high global warming potential (GWP) refrigerants. CO_2-based refrigerants have emerged as a promising choice in response, distinguished by its low GWP and reduced flammability, compared to formulated hydrofluoroolefins, thus offering a safer and sustainable solution in the context of next generation drop-in refrigerants. This study presents a machine-learning-based methodology to estimate the saturation properties of CO_2-based mixtures, allowing for the precise tuning of molecular-based models like the polar soft-SAFT, used for technical evaluation, without relying on experimental data, often unavailable for such systems. The approach departs from the thermodynamic characterization of several pure-components, including novel fluorine-based refrigerants. The parametrization allows an excellent description of the vapor pressure, saturated densities, and latent heat. Next, a constant, temperature-independent binary parameter is used to estimate the solubility profiles of CO_2-derived mixtures in selected refrigerants. The model effectively captures azeotropic and zeotropic behaviors, demonstrating its strength in fine-tuning solubility with minimal corrections. Subsequently, data from the molecular characterization via polar soft-SAFT is used as output targets to train a machine learning algorithm based on artificial neural networks, enabling the prediction of mixture saturation properties out of the training dataset's scope. Using COSMO σ-profiles, the developed ANN demonstrates high efficiency in predicting saturation bubble and dew temperatures, achieving $R^2 > 0.9999$, RMSE< 0.0959, AARD < 0.0220 %, and NMAD of 0.00044. Statistical analysis confirms minimal mean deviations, with outliers limited to 2.63 % for bubble and 2.44% for dew phase predictions, respectively.

1. Introduction

Refrigeration and air conditioning technologies have experienced a transformative evolution [1], driven by the changing landscape of fluorinated gases (F-gases), a subgroup of greenhouse gases (GHGs) highly efficient in refrigeration cycles. Unfortunately, these gases have high global warming potential (GWP), and hence, they have played a crucial role in altering Earth's climate, with hydrofluorocarbons (HFCs) [2,3] being a notable example in this respect.

The evolution of refrigerants traces a path through various generations [4], leading to the adoption of simple yet functional compounds as with water, ammonia, carbon dioxide (CO_2), and methyl chloride (CH_3Cl) in early applications [5]. Initially revolutionary [6], these first-generation refrigerants were eventually replaced due to safety concerns including ammonia's high toxicity, CO_2's high pressure requirements, and CH_3Cl's flammable and toxic nature. This quest for safer alternatives [7] paved the way for the development of new generations of refrigerants [8,9]. The transition started with chlorofluorocarbons (CFCs) as with Freon (R12) in the 1930s, valued for non-flammability, but later banned in the Montreal Protocol because of its effect on the ozone layer hole. Next, hydrochlorofluorocarbons (HCFCs) such as chlorodifluoromethane (R22) emerged in the 1980s, with lesser ozone impact, but still highly pollutant in terms of GWP. Finally, hydrofluorocarbons (HFCs) as with 1,1,1,2-tetrafluoroethane (R134a) were proposed and effectively used in the 1990s; however, while ozone-safe, they are currently targeted to be replaced under the Kigali Amendment [10] due to their high-GWP.

Each stage of this evolution, defined by distinct environmental impacts and regulatory measures [11], has propelled the advancement of

* Corresponding authors.
E-mail addresses: lourdes.vega@ku.ac.ae (L.F. Vega), felix.llovell@urv.cat (F. Llovell).

https://doi.org/10.1016/j.jcou.2025.103072
Received 7 January 2025; Received in revised form 4 March 2025; Accepted 21 March 2025
Available online 4 April 2025

Abbreviations

1	Null Flammability
2L	Mid-Flammability
3	High Flammability
A	Null Toxicity
A^{assoc}	Association Helmholtz Energy Density
A^{chain}	Chain Formation Helmholtz Energy Density
A^{polar}	Polar Term Helmholtz Energy Density
A^{ref}	Reference Fluid Helmholtz Energy Density
A^{res}	Residual Helmholtz Energy Density
AI	Artificial Intelligence
AAD	Absolute Average Deviation
AARD	Absolute Average Relative Deviation
Act	Actual Data
ANN	Artificial Neural Networks
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
B	High Toxicity
b	Bias
C2	Compound Number 2
CAS	Chemical Abstracts Service
CCUS	Carbon Capture, Utilization and Storage
CFCs	Chlorofluorocarbons
COSMO-RS	Conductor like Screening Model for Realistic Solvents
def-TZVP	Double-Electron, Triple-Zeta Valence with Polarization
DFT	Density Functional Theory
e	Elementary Charge
EoS	Equation of State
EU	European Union
F	Fluorinated
GHG	Greenhouse Gas
GWP	Global Warming Potential
HCs	HydroCarbons
HCFOs	Hydrochlorofluoroolefins
HFCs	Hydrofluorocarbons
HFOs	Hydrofluoroolefins
HLX	Hidden Layer X
HOs	Hydroolefins or Alkenes
I	Input
KPI	Key Performance Indicator
LB	Lorentz-Berthelot
LJ	Lennard-Jones
m_i	Chain Length
MSE	Mean Squared Error
N	Number of Observations
NIST	National Institute of Standards and Technology
$NMAD_w$	Weighted Normalized Median Absolute Deviation
O	Output
P^i	Vapor Pressure of Molecule i
P^M	Vapor Pressure of Mixture M
Pred	Predicted Data
QSPR	Quantitative Structure Property Relationships
R^2	Coefficient of Determination
RE	Refrigeration Effect
RMSE	Root Mean Square Error
S_σ^i	Surface Charge Density for Molecule i
S_σ^M	Surface Charge Density for Mixture M
SAFT	Statistical Associating Fluid Theory
SCF	Self-Consistent Field
SD_{av}	Average Standard Deviation
SDR	Standardized Residuals
sHCs	Saturated Hydrocarbons or Alkanes
SMILES	Simplified Molecular Input Line Entry System
Tansig	Hyperbolic Tangent
T_b	Bubble Temperature
T_d	Dew Temperature
VLE	Vapor-Liquid Equilibria
var	Variance
w	Weight
x_i	Mole Fraction of Compound i in the Liquid Phase
x_p	Fraction of Polar Segments within the Molecule
y_i	Mole Fraction of Compound i in the Vapor Phase

Formulas, Units, and Greek Symbols

Å	Ångström
ξ_{ij}	Energy Binary Parameter in the polar soft-SAFT Equation of State
μ_i	Experimental Dipole Moment in Vacuum
Q_i	Experimental Quadrupole Moment in Vacuum
ε_i	Dispersive Energy between Segments
# N_{HL}	Number of Neurons in each Hidden Layer
ρ^i	Saturated Density of Molecule-type i
η_{ij}	Size Binary Parameter
σ	Specific Charge Density
σ_i	Specific Diameter

sustainable cooling technologies in terms of refrigerant development, furthering global initiatives to adopt more environmentally friendly solutions. The American Innovation and Manufacturing Act of 2020 (AIM Act) [12], is directing a substantial reduction in HFCs, targeting a 40 % decrease by 2024 and 85 % by 2036. The Environmental Protection Agency's (EPA) phasedown plan of HFCs further restricts the manufacture, import, and sale of non-compliant products with GWP higher than 700 in new air-conditioning and refrigeration equipment. Simultaneously, the EU's F-gas regulation, effective since 2015 [13], strives for a 79 % HFC reduction by 2030, with a recent proposal for an even more ambitious phase-down, thus aligning with the European Green Deal and the European Climate Law. Such international regulatory efforts are not only responding to environmental imperatives, but are actively shaping the landscape of refrigerant technology, promoting the emergence of innovative, low-GWP formulas, as well as improving the efficiency of cooling systems.

At this stage, the deployment of pure working fluids remains limited by the availability of suitable compounds having all desirable properties [14,15], leading the direction towards the formulation of refrigerant blends from at least two working fluids, capable of meeting the desired technical, environmental, and safety requirements. Provided in Table 1 is a summary of available commercial blends and their individual components. The performance of the blend depends mainly on the choice of individual components and their ratios. Predominantly, HFC difluoromethane (R32) is currently at the forefront [16], emerging as the most popular choice for deployment in low-GWP refrigerant blends through its GWP = 677, closely followed by the ultra-low-GWP hydrofluoroolefin 2,3,3,3-tetrafluoropropene (R1234yf), and saturated hydrocarbon propane (R290) with GWP ~ 3 [17,18], as clearly identified in Table 1. Refrigerants including 1,1-difluoroethane (R152a), trans-1,3,3,3- tetrafluoroprop-1-ene (R1234ze(E)), and isobutane (R600a) also find usage in air conditioning applications [19–21], albeit to a lesser extent, while once-popular agents R134a and 1,1,1,2,3,3,3-heptafluoropropane (R227ea), are currently experiencing a downturn in adoption,

Table 1
Overview of current commercial low/medium-GWP refrigerant blends in the market, including their GWP, ASHRAE safety classifications, and mass composition percentages.

	GWP	ASHRAE	R32	R1234yf	R290	R152a	R1234ze(E)	R600a	R134a	R1270	RE170	R227ea	R600	R125	CO_2	R170	R1132a	R1336mzz(E)	R1336mzz(Z)	R13I1	R1130t
R436C	1	A3			95			5													
R510A	1	A3						12			88										
R432A	2	A3								80	20										
R514A	2	B1																	75		25
R443A	3	A3			40			5		55											
R433A	3	A3			70					30											
R433C	3	A3			75					25											
R433B	3	A3			95					5											
R436A	3	A3			56			44													
R436B	3	A3			52			48													
R511A	3	A3			95						5										
R441A	3	A3			55			6					36			3					
R429A	14	A3				10		30			60										
R435A	26	A3				20					80										
R431A	38	A3			71	29															
R444A	93	A2L	12			5	83														
R430A	95	A3				76		24													
R516A	131	A2L		77		14			9												
R451A	133	A2L		90					10												
R445A	135	A2L					85		9						6						
R457A	139	A2L	18	70		12															
R440A	140	A2			1	97			2												
R466A	143	A2	49											11						40	
R465A	143	A2	21	71	8																
R459B	143	A2L	21	69			10														
R455A	145	A2L	22	75											3						
R454C	148	A2L	21	79																	
R471A	148	A1					79					4						17			
R468A	148	A2L	22	75													3				
R451B	160	A2L		89					11												
R512A	189	A2				95			5												
R454A	238	A2L	35	65																	
R449A	238	A2L	24				25		26					25							
R457B	251	A2L	35	55		10															
R444B	295	A2L	42			10	48														
R515B	299	A1					91					9									
R515A	387	A1					88					12									
R459A	460	A2L	68	26			6														
R454B	466	A2L	69	31																	
R446A	470	A2L	29				68						3								

due to their higher GWP, with 1300 and 3350, respectively [22].

Newly introduced refrigerants [23], aligning with the latest EU regulations [13], such as dimethylether (RE170), 1,1-difluoroethylene (R1132a), trans/cis-1,1,1,4,4,4-hexafluoro-2-butene (R1336mzz(E/Z)), trifluoroiodomethane (R13I1), and notably CO_2, are beginning to surface as viable low-GWP options (*i.e.,* GWP = 1 – 15), although their presence remains limited in refrigerant blends. Particularly, CO_2 is distinguished by its potential as a working fluid [24–26], owed to its non-flammability and environmental friendliness in terms of GWP, despite facing technical challenges arising from its high critical pressure and low critical temperature, limiting its range of applications compared to standard F-gases. This limitation typically caps the CO_2 content in commercial blends to 3 – 6 % as seen with R445A and R455A [27]. However, this is currently under reconsideration by industry professionals tipping the scale towards environmental benefits over safety standards [28,29], in response to evolving directives set by governmental authorities. In addition, it is notable that most commercial blends, highlighted in Table 1 [30,31], with a GWP < 150 are characterized by high flammability (2/3), toxicity (B), or glide (T_G >10 K), with noteworthy exceptions including R516A (A2L), R451A (A2L), or R471A (A1), which are emerging as potential replacements to the industry benchmark R134a.

As seen from Table 1, the presence of CO_2 in commercial refrigerant blends has been limited during the last 40 years. Blending CO_2 with HFOs or low-GWP HFCs in transcritical systems has been found to reduce pressure standards and improve COP, though increasing flammability and thermal instability [32,33]. Therefore, there is a clear opportunity to delve deeper into the rational design of CO_2-based blends with a wider array of refrigerants, such as HFCs, HFOs, HCFOs, and even hydrocarbons (HCs), targeting improved working fluid properties. This scenario reveals a substantial scope for innovative development in refrigerant technology, particularly emphasizing the role of CO_2, which could be a game-changer in the pursuit of environmentally sustainable cooling solutions. Interestingly, CO_2-based mixtures with refrigerants have undergone extensive study, with their saturation properties for common HFCs like R32, pentafluoroethane (R125), R134a, R152a, R227ea, trifluoromethane (R23), and fluoromethane (R41) well-documented in the NIST database [34]. The database's scope was later extended to cover mixtures with saturated hydrocarbons and hydroolefins including ethane (R170), R290, the butane-isopentane (R600-R601a) sequence, ethylene (R1150), propylene (R1270), along with RE170, R1234yf, R1234ze(E), and the novel R13I1. Advancements in this field have led to the generation of new saturation data for CO_2 mixtures with novel cooling agents, including trifluoroethylene (R1123) [35], 3,3,3,-trifluoropropene (R1243zf) [36], 1-ethoxy-1,1,2,2,3,3,4,4,4-nonafluorobutane (R7200) [37], cis-1,3,3,3-tetrafluoropropene (R1234ze(Z)) [38], R1336mzz(E) [38], and trans-1-chloro-3,3,3-trifluoropropene (R1233zd(E)) [39] from 2019 to 2025. Key additional properties including single-phase density [40], viscosity [41], and thermal conductivity [42] have also been assessed for CO_2-based refrigerant mixtures with selected alkanes, alkenes, R1234yf, R32, R134a, and R125. Data on density is also available for CO_2 blends with refrigerants like R1234ze(E), R152a, and R41, enhancing the understanding of their intrinsic behavior in a range of applications [40].

Given the considerable amount of available data, there is a chance to implement and validate accurate models to describe the thermodynamic behavior of CO_2-based refrigerant blends, either to evaluate their performance in a given refrigeration cycle or extrapolate to predict the thermodynamic properties of new and untested CO_2-based refrigerant blends. Among the approaches available, Yang *et al.* [43] applied entropy scaling in conjunction with the REFPROP [34] reference equation of state to model the thermal conductivity of various pure and mixed refrigerants. In another work, Bell *et al* [40]. recognized the applicability of the NIST REFPROP semi-empirical model [34] for engineering applications; however, concerns were raised regarding its predictive accuracy, the transferability of parameters to new mixtures, and the extensive experimental data required for developing new blends. These limitations, however, can be addressed by the use of physics-rooted molecular-based equations of state (EoS), such as those from the statistical associating fluid theory (SAFT)-family, which have been instrumental in accurately modeling a wide range of thermophysical properties for complex systems, including natural gas mixtures with impurities [44] hydrogen [45], hybrid solvents [46], glycols [47], ionic liquids [48] and low-GWP refrigerants [15,49].

Even though accurate SAFT-based models have demonstrated highly predictive ability for modeling pure fluids, the extension to mixtures requires the use of binary parameters regressed to phase equilibria data, decreasing the capacity of extrapolation to other systems and conditions with limited or unavailable experimental data. Advanced thermodynamic models are required due to the non-ideality of mixtures with CO_2, arising from the asymmetrical energy and size between the quadrupolar and linear CO_2 molecules and other components in CO_2-based refrigerant blends.

In this regard, current scientific efforts aim to cover these gaps in extrapolation capacity by coupling artificial intelligence (AI) methods with molecular-based equations. In particular, the integration of the polar soft-SAFT EoS with AI-based approaches provides a holistic methodology only dependent on Conductor-like Screening Model for Realistic Solvation (COSMO-RS) molecular descriptors as input parameters, enabling rapid and cost-effective first-screening predictions without the need of experimental efforts. Their molecular-level resolution captures dipolar and quadrupolar moments, hydrogen bonding, and dispersion forces, enabling soft-SAFT to accurately represent polar and associative effects in CO_2-based fluids and other F-based refrigerants with strong electrostatic interactions. In the past, a wide-range of properties ranging from viscosity and pH to flammability, surface tension, and conductivity have already been estimated through the use of quantum chemistry descriptors applied to ionic liquids (ILs) [50–52], deep eutectic solvents (DESs) [52–57], eutectic solvents (ESs) [58], ester-alkanes [59], fluorinated refrigerants [60], and polymers [61]. In recent years, some studies have started incorporating COSMO-RS and other AI-based models to predict CO_2 solubility in DESs. For instance, a recent study [62] reported an AARD of 2.47 %–21.90 % for a QSPR model applied to various DES systems, significantly outperforming COSMO-RS, which exhibited deviations of 42.91 %–88.17 %. Additionally, Chen *et al.* [63] have explored solid-liquid-gas equilibrium predictions for CO_2 + organic compound mixtures using COSMO-SAC-based approaches. When combined with the Peng-Robinson EoS, COSMO-SAC resulted in deviations of 6.25 K in temperature, 0.071 in liquid mole fraction, and 21 % in liquid molar volume. Furthermore, absolute average relative deviations in the liquid molar volume of several CO_2-based binary mixtures remained high, around 20 %–30 %. Focused on refrigeration, Deng *et al.* [64] combined artificial neural networks with genetic algorithms (GA) to model normal boiling points, achieving an absolute average deviation of 1.87 %. Similarly, a recent approach by Liu *et al.* [65] proposed an extreme learning machine (ELM) and an ensemble decision tree boosted algorithm (EDT Boosted) to predict the normal boiling point, using over a dozen different molecular groups and a topological index. Following similar patterns, a neural network-based approach was developed by Wang *et al.* [66] on predicting the critical temperature of binary refrigerant mixtures, employing a graph neural network (GNN) while outperforming conventional empirical and equation-of-state-based methods. However, these models remain purely empirical and limited in terms of functionality, with one model descriptive of a single thermodynamic property, thus requiring multiple models and expansive databases to cover a wide range of properties and fluids. In light of new data, the model can be retrained to enhance its accuracy and computation range. The coupling of AI-based approaches with molecular-based EoSs can be a rewarding approach combining large data processing with physics-oriented modeling.

Despite advances in AI-based property prediction, a gap remains in

accurately predicting saturation properties of refrigerant systems through data-driven models not dependent on empirical data—an approach not yet applied. While existing models have predominantly focused on pure components or isolated thermophysical properties, the application of QSPR techniques based on molecular descriptors and novel EoS for predicting saturation properties in both liquid and vapor phases remain limitedly explored. In this direction, a preliminary contribution from Alkhatib *et al.* [67] applied ANN models trained on soft-SAFT molecular parameters to predict the saturation properties of pure refrigerants, further extending the use of machine learning in thermodynamic modeling. Now, the current study aims to progress on this field by integrating the polar soft-SAFT EoS [68], with an AI-based technique, namely, artificial neural network (ANN) to predict dew and bubble points of CO_2-based refrigerant blends, thus addressing a significant need in the field of refrigeration research. This is done employing input molecular-based descriptors obtained from the quantum chemistry model, Conductor-like Screening Model for Realistic Solvation, to predict VLE data for CO_2-based refrigerant blends. As mentioned, COSMO-RS descriptors effectively capture factors influencing phase saturation properties by considering surface charge density and molecular interactions, including the presence of specific functional groups, polarity, oxygen content, hydrogen bonding capacity, and the presence of heteroatoms like fluorine, chlorine or iodine. This is necessary to reliably predict VLE data required for calibrating molecular-based EoSs, such as polar soft-SAFT, to accurately describe the behavior of CO_2-based refrigerants including thermophysical, transport, and energetic properties, required for the technical evaluation of these blends for specific cooling systems.

2. Methodology

Highlighted in Fig. 1 is the detailed flowchart methodology adopted in this work, outlining the step-by-step process in the machine learning technique. The polar soft-SAFT EoS is used to model the VLE behavior of selected CO_2-based refrigerant mixtures, with the experimental dew and bubble point temperatures used as output data for ANN training, thus developing two independent codes for each response. Both ANN-based models adhere to the same setup conditions, with the optimal configuration of neurons in both hidden layers dependent on each output's main outcomes. Note that the polar soft-SAFT EoS for binary CO_2-mixtures is calibrated with available experimental data to ensure accurate representation of these systems. The choice of using polar soft-SAFT computed VLE data rather than directly using the experimental data is to ensure finite discretization of VLE data at the entire span of mole composition, required to ensure reliable ANN training. The COSMO-RS based molecular descriptors are used as ANN inputs, descriptive of CO_2 and other components in the binary blends. With the required training dataset, the ANN is developed with a fixed structure in terms of number of hidden layers and activation functions, yet, allowing number of neurons in the hidden layer to be optimized for higher model accuracy. The performance of the model is rigorously assessed using a wide set of statistical-based key performance indicators (KPIs). After validation and testing, the ANN model is used to predict the VLE behavior of systems without experimental data, which can be used to fine-tune the polar soft-SAFT models, allowing the holistic assessment of these new blends.

2.1. Polar soft-SAFT EoS

The polar soft-SAFT EoS (see Eq. (1)), is a variant of the original SAFT EoS [69], based on Wertheim's first-order thermodynamic perturbation theory (TPT1) for associating fluids [70–73]. The polar soft-SAFT version, written in terms of the residual Helmholtz energy, enables explicit modeling of fluids with multi-polar interactions, such as dipole and quadrupole, extending the soft-SAFT equation of 1997 [74] post-2020 [68]. The model computes the residual Helmholtz energy as the sum of distinct microscopic terms (*i.e.,* reference, chain, association, and polar), effectively capturing the complexity of molecular interactions and its intrinsic effects. This approach is particularly effective in addressing complex phase behavior and thermodynamic properties, including enthalpy and entropy.

$$A^{res} = A^{ref} + A^{chain} + A^{assoc} + A^{polar} \tag{1}$$

The reference term, denoted as A^{ref}, incorporates segment-segment interactions within the fluid. This is modeled in soft-SAFT using a Lennard-Jones (LJ) intermolecular potential based on the LJ EoS of Johnson *et al.* [75], which accounts for both repulsive and attractive forces in a single formulation representative of the fluid. The chain term, A^{chain}, addresses the contributions arising from the connectivity of individual segments, forming chain-like structures, while A^{assoc}, is crucial for modeling directional and strong-range interactions such as hydrogen bonding [76,77]. Finally, the polar term, A^{polar}, explicitly considers multipolar interactions like dipolar and quadrupolar forces, incorporating the pair and triplet integral correlation functions for the reference fluid, from Luckas *et al* [78]. The dynamics between pairs and trios of multipolar molecules are formulated on the roots of Gubbins and Twu's theories [79,80], adapted to chain-like polar fluids by Jog *et al.*'s segment-approach [81,82] and further refined according to approximations of the type of Padé [83] for enhanced precision. For detailed insights on the integration and application of this term into soft-SAFT,

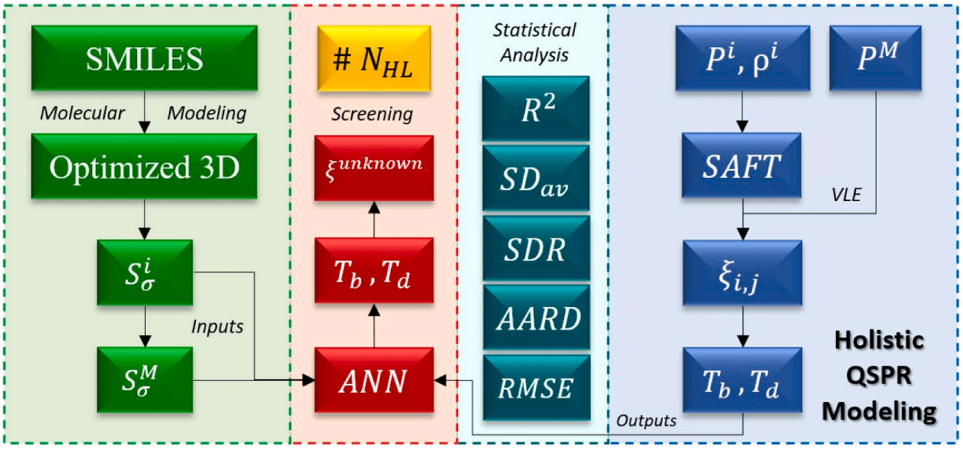

Fig. 1. Flowchart and integrated modeling framework used in the QSPR modeling. Green is used for COSMO-RS molecular modeling, red for machine learning and subsequent screening of unknown mixtures, blue for polar soft-SAFT characterizations, and turquoise for model evaluation and external validation according to selected statistical Key Performance Indicators (KPIs). The reader is referred to the nomenclature for a full understanding of the labels of each box.

the reader is referred to previous literature [44,84–86].

In this work, refrigerant molecules are modelled as non-associating and chain-like fluids, which are either dispersive or polar, depending on the molecular structure and energy profile [15,49,67,87]. The key polar soft-SAFT molecular parameters applicable to all types of molecules, include chain length (m_i), LJ segment diameter (σ_i), and dispersive van der Waals energy among monomers (ε_i), are regressed using available pure refrigerants saturated liquid density and vapor pressure data.

In the case of polar fluids (such as the fluorinated refrigerants and CO_2), additional parameters include dipole (μ_i) and/or quadrupole (Q_i) moments, influenced by halogen-induced asymmetrical charge distributions, which are set based on their experimental value in vacuum. In this context, halogens in refrigerants exhibit asymmetrical charge distributions and, as a consequence, distinct dipole effects indicative of strong polar behavior, as appreciated in Table 2, while CO_2, with its linear structure, results in a notable quadrupole moment. Consequently, dipole and quadrupole effects are assigned to chain spheres in its coarse-grain representations, ensuring accurate modeling of CO_2-based mixtures with crossed multipolar interactions. Also, molecules like

Table 2
CAS Number, SMILES, and dipole moment coarse-grain (CG) representation for all compounds assessed in this work. Atoms are assigned as follows: C as grey, H as white, O as red, F as reddish, Cl as green, and I as purple. The position and length of the arrow is indicative of the area and value of the dipole moment.

Compound	CAS	SMILES	Dipole	Compound	CAS	SMILES	Dipole
R41	593–53–3	CF		R1336mzz(Z)	692–49–9	C(=CC(F)(F)F)C(F)(F)F	
R32	75–10–5	C(F)F		R1336yf	374–39–0	C=C(C(F)(F)F)C(F)(F)F	
R23	75–46–7	C(F)(F)F		R1345fz	374–27–6	C=C(C(F)(F)F)C(F)(F)F	
R161	353–36–6	CCF		R1233zd(E)	102687–65–0	C(=CCl)C(F)(F)F	
R152a	75–37–6	CC(F)F		R1224yd(E)	111512–52–8	C(\C(F)(F)F)=C(/F)\Cl	
R134a	811–97–2	C(C(F)(F)F)F		R1224yd(Z)	111512–60–8	C(\C(F)(F)F)=C(\Cl)F	
R125	354–33–6	C(C(F)(F)F)(F)F		R170	74–84–0	CC	
R245fa	460–73–1	C(C(F)F)C(F)(F)F		R290	74–98–6	CCC	
R236fa	690–39–1	C(C(F)(F)F)C(F)(F)F		R600	106–97–8	CCCC	
R227ea	431–89–0	C(C(F)(F)F)(C(F)(F)F)F		R600a	75–28–5	CC(C)C	
R1123	359–11–5	CF=C(F)F		R601	109–66–0	CCCCC	
R1132a	75–38–7	C=C(F)F		R601a	78–78–4	CC(C)CC	
R1132(E)	1630–78–0	C(=CF)F		R1150	74–85–1	C=C	
R1225ye(Z)	5528–43–8	C(=C(C(F)(F)F)F)F		R1270	115–07–1	C=CC	
R1234yf	754–12–1	C=C(C(F)(F)F)F		1-butene	106–98–9	C=CCC	
R1234ze(E)	29118–24–9	C(=CF)C(F)(F)F		CO_2	124–38–9	C(=O)=O	
R1234ze(Z)	29118–25–0	C(=CF)C(F)(F)F		R13I1	2314–97–8	C(F)(F)(F)I	
R1243zf	677–21–4	C=CC(F)(F)F		RE170	115–10–6	COC	
R1336mzz(E)	66711–86–2	C(=CC(F)(F)F)C(F)(F)F		R7200	163702–05–4	CCOC(C(C(C(F)(F)F)(F)F)(F)F)(F)F	

hydrocarbons and certain HFOs, including R1132(E) and R1336mzz(E), exhibit minimal dipoles due to their more symmetrical electronic distribution and/or absence of electronegative atoms creating charge gradients across the molecular structure. An additional parameter for polar fluids is the fraction of polar segments (x_p), which is pre-set based on physical principles and consistency across related chemical families and structurally similar compounds. This procedure streamlines the fitting process, ensuring a reduced number of adjustable parameters, while enhancing the model's predictive capability, demonstrated from earlier contributions [15,44,49,85].

The application of the model to multicomponent mixtures is done using the van der Waals one-fluid theory (vdW1) [88], used to extend the reference term to account for variations in molecular volume and energy. This is applied using the modified Lorentz-Berthelot (LB) combining rules to compute crossed-size and energy for the resulting conformal pseudo-binary fluid corrected by binary parameters η_{ij} and ξ_{ij}, respectively.

$$\sigma_{ij} = \eta_{ij} \cdot \left(\frac{\sigma_{ii} + \sigma_{jj}}{2} \right) \tag{2}$$

$$\varepsilon_{ij} = \xi_{ij} \cdot \left(\varepsilon_{ii} \cdot \varepsilon_{jj} \right)^{1/2} \tag{3}$$

The model is used in a predictive capacity when adjustable binary parameters in Eqs. (2) and (3) are set to one. Otherwise, and as a standard procedure, these parameters are fine-tuned using experimental phase equilibrium binary data such as VLE data, enabling soft-SAFT to effectively regress mixture behaviors quantitatively. As an example, the interactions between oxygen atoms in CO_2 and hydrogen groups in refrigerants are accounted by the dispersion term, coupled with the energy binary parameter ξ_{ij}, addressing the mixtures' non-ideal behavior. None of the mixtures of this study has required the implementation of the η_{ij} size binary parameter correction, in line with prior studies using polar soft-SAFT [48]. In cases where two compounds exhibit areas of immiscibility, a second binary parameter correction is sometimes applied to improve the accuracy of phase equilibrium predictions, as demonstrated by Albà *et al.* [85].

2.2. AI-ANN for modeling saturation properties of CO_2-based refrigerant blends

2.2.1. Output training data assembly

In the context of data assembly and preprocessing, we have compiled a wide-ranging array of data, serving as outputs to the AI-model, consisting of 531 data points for vapor-liquid equilibria (VLE) of predominantly CO_2-based refrigerant blends. Specifically, this collection features 494 data points covering CO_2 compositions (both in liquid and vapor phases) ranging from 5 % to 95 % in mole fraction, combined with a total of 26 refrigerants of different nature. These blends, for which experimental VLE data is available in the literature, are categorized as follows: 8 HFCs including R41, R32, R23, R152a, R125, R227ea [34], R161 [89], and R134a [90]; 6 HFOs, namely R1123 [35], R1243zf [36], R1234yf [91], R1234ze(E) [92], R1234ze(Z) [38], and R1336mzz(E) [38]; 1 HCFO, R1233zd(E) [39]; 6 sHc comprising R170, R290, R600, R600a, R601, and R601a [34]; 3 HOs — R1150, R1270, and 1-butene [34]; along with R13I1 [93], and RE170 [34]. Additionally, the output database covers saturation targets for all 27-constituent pure-components, supplemented by 10 additional F-based single refrigerants. These include R245fa, R236fa, R1225ye(Z), R1336mzz(Z), R1132(E), R1132a, R1345fz, R1336yf, R1224yd(E), and R1224yd(Z). Their inclusion, despite lacking binary VLE data with CO_2, broadens the scope of this analysis. R7200 and R1130(E) were excluded from the model's database due to extreme saturation divergence from the threshold established by its counterparts, posing a risk of inaccurate predictions in CO_2-based mixture modeling, particularly for medium-low temperature applications. Through our extensive coverage of diverse compounds and compositions, it is aimed at enhancing the accuracy, applicability and reliability of the presented AI-based model, ensuring solid and accurate predictions for a wide range of CO_2-based refrigerant mixtures. The dataset sufficiently covers the target feature range for low-temperature applications (220 - 300 K), with limited data above 330 K due to CO_2's low critical temperature.

2.2.2. Input descriptors

In the pursuit of understanding the intricate dynamics present in CO_2-based refrigerant blends, a computational methodology based on quantum chemistry is applied, with COSMO-RS model used to provide molecular-level descriptors, thereby establishing a link between molecular features and their macroscopic behavior. The method involves first converting the Simplified Molecular Input Line Entry Specification (SMILES) codes, into 3D representations using Turbomole software and its TmoleX v4.5.1 interface. The molecular geometry is subjected to optimization at the Density Functional Theory (DFT) level using the def-TZVP basis set and the Becke–Perdew 86 (BP86) generalized gradient approximation [58], integral to the DFT calculation sequence for σ-profile generation. The Self-Consistent Field (SCF) convergence criterion is set at a threshold of 1×10^{-6} Hartree for enhanced precision in the calculations.

Post optimization, the molecular details are exported as COSMO [94] files and used as input data to generate the σ-profiles [95] in COSMO-ThermX software (version 19.0.5) through FastSigma coding, providing a comprehensive analysis of molecular characteristics and interactions in the ground-state level. σ-profiles, spanning a range of $\pm$ 0.03 e/Å^2 and composed of 61 discretized data points [96], present a thorough insight into molecular surfaces, effectively delineating between polar and nonpolar areas of influence, effectively mapping the electron density on a molecule's surface. This is fundamental to gaining detailed insights into the interactions and ground-state geometric configurations of molecules in various environments [97,98].

The additive nature of σ-profiles [52,99,100], derived from individual atomic contributions, serves as empirical descriptors for accurately modeling refrigerant mixtures, as shown in Eq. (4). The computation of a mixture's σ-profile is a linear combination of its individual components, proportionally weighted by their mole fractions in the mixture. The 61 discretized σ-profiles for CO_2-based refrigerant blends serve as inputs to the ANN model. This approach proves particularly effective in the simulation of complex, multi-component mixtures of molecular liquids and derivatives [101], providing a versatile, accessible and user-friendly approach for the rapid screening of thermophysical properties of new mixture candidates.

$$S_\sigma^M = \sum_{i=1}^{2} (x_i) \cdot \left(S_\sigma^i \right) = (x_{CO_2}) \cdot \left(S_\sigma^{CO_2} \right) + (1 - x_{CO_2}) \cdot \left(S_\sigma^{C2} \right) \tag{4}$$

2.2.3. ANN structure and details

ANNs [102], chosen for their ability in processing complex, nonlinear data, are particularly suited for analyzing detailed molecular fluid behaviors. Structurally analogous to the human brain's network of neurons, ANNs consist of interconnected layers of nodes: input, hidden, and output layers, whose their interconnected weights and biases undergo adaptive learning, marked by iterative training and systematic evaluation against targeted outcomes [103,104]. This capability renders ANNs highly effective for pattern recognition and predictive modeling.

A feed forward ANN is employed in this work, with a dual hidden layer structure, employing the Levenberg-Marquardt ('trainlm') algorithm for training [105–107], known for its superior efficiency in medium-sized datasets, as it combines both gradient descent ('traingd') and Gauss-Newton ('traincg') methods. While the Levenberg-Marquardt algorithm fits this study's demands in high-dimensional, nonlinear data modeling, its convergence algorithm makes it less suitable for larger-scale scenarios; instead, Broyden-Fletcher-Goldfarb-Shanno (BFGS) ('trainbfg') is advised, requiring higher memory consumption,

while maintaining an appropriate balance between efficiency, seed, and resource utilization. The hyperbolic tangent ('tansig') is selected as the activation function of the ANN's hidden layers [52,108], favored for its efficient backpropagation, zero-centered output range, and effective management of gradient flows *via* partial derivatives. The ANN is pre-designed with a fixed structure, specifying the number of hidden layers (two) and activation functions, based on the results from previous contributions [60,67] where various network architectures, input dimensions and activation functions were systematically evaluated to ensure optimal predictive performance of diverse refrigerant properties, such as flammability and density. The neurons are fine-tuned in both intermediate layers (HL1/ HL2) to enhance accuracy [109].

The *tanh* function has been chosen to yield the output *O* for each neuron *j* in the hidden layers, as shown in Eq. (5), with this output then serving as input for the subsequent layer in the ANN, accommodating *n* inputs per neuron. This aids in optimizing weight adjustments and loss minimization, as a result of addressing the vanishing gradient effect, a key aspect in training deep neural models within MATLAB's computational environment. *Tanh*'s centered output scope of ±1 enhances training convergence by ensuring balanced neuron activation, stabilizing gradients, and improving the efficiency and stability of the network. Subsequently, the output is computed using the linear 'purelin' activation function, contributing to a direct, proportional range of continuous output values. *Purelin* is particularly suited for regression tasks spanning a broad range of applications, fitting scenarios requiring a linear input-output relationship.

$$O_j = \tanh\left(\sum_{i=1}^{n}\left(\omega_{ij}\cdot S_\sigma^i\right)+b_j\right) = \frac{e^{\sum_{i=1}^{n}\left(\omega_{ij}\cdot S_\sigma^i\right)+b_j} - e^{-\sum_{i=1}^{n}\left(\omega_{ij}\cdot S_\sigma^i\right)+b_j}}{e^{\sum_{i=1}^{n}\left(\omega_{ij}\cdot S_\sigma^i\right)+b_j} + e^{-\sum_{i=1}^{n}\left(\omega_{ij}\cdot S_\sigma^i\right)+b_j}} = \frac{e^{2\cdot\sum_{i=1}^{n}\left(\omega_{ij}\cdot S_\sigma^i\right)+b_j} - 1}{e^{2\cdot\sum_{i=1}^{n}\left(\omega_{ij}\cdot S_\sigma^i\right)+b_j} + 1} \quad (5)$$

Post-fitting residual analysis is necessary to verify the model's linear distribution and continuous assumption, reinforcing its validity. This setup balances *tanh*'s non-linear processing with *purelin*'s linear treatment to produce unrestricted, continuous output predictions, enhancing the network's predictive potential. Building on the ANN's robust framework described, the ANN model processes input data derived from σ-profile descriptors, adopting Mean Squared Error (MSE) as the loss function for training, selected for its straightforward interpretability and sensitivity to large errors, with Mean Absolute Error (MAE) or Huber loss as alternatives for outlier-rich datasets. The training process includes early stopping criteria to avert overfitting, while simultaneously applying a data partitioning strategy to enhance the model's generalizability. Key training parameters, such as epochs and data division, are carefully tuned to ensure the model's accuracy in predicting molecular fluid characteristics. '*dividerand*' is the default function used for data partitioning in MATLAB, randomly allocating data into training, validation, and testing subsets. The '*net.divideParam*' property, where '*trainRatio*', '*valRatio*', and '*testRatio*' are set to 80 %, 10 %, and 10 % respectively, matches the criteria specified in earlier discussions [110]. The training set is used for computing the gradient and updating the network's weights and biases, set to operate for a maximum of 1000 epochs and targeting a MSE of 10^{-20}. During training, a rise in validation error for six consecutive epochs triggers MATLAB's default early stopping due to overfitting, halting the training and reverting the network to its lowest validation error state.

2.2.4. Statistical analysis for evaluating ANN model performance

The performance of the developed multi-layer ANN model is evaluated using a complete set of statistical metrics, including the Coefficient of Determination (R^2) for correlation adequacy, Root-Mean-Square Error (RMSE) for predictive accuracy, Average Absolute Relative Deviation (AARD) for precision, and Average Standard Deviation (SD_{av}) for data variability and dispersion, alongside Normalized Median Absolute Deviation (NMAD), pivotal for assessing multi-property scenarios, providing a comprehensive insight into the model's fitting and predictive capabilities. Notably, higher R^2 values are preferred, while 'lower the better' criteria applies to the remaining KPIs for enhanced performance. This attribute complements the first quartet of metrics, all defined in Eqs. (6) - (9), which distinctly assess the model's performance in both target responses. In contrast, NMAD, derived from the expression outlined in Eq. (10), delivers a wide-ranging outline across the entirety of output runs, normalizing deviations in effect and diminishing its sensitivity to outliers in comparison with RMSE. R^2 measures the model's predictive accuracy, RMSE captures overall error magnitude, and AARD assesses the average relative discrepancy between predictions and experimental data. SD_{av} ensures robustness by quantifying variability, while NMAD normalizes deviations for better cross-dataset comparison. For detailed information on the formulations and terminologies of Eqs. (6) - (10), readers are encouraged to refer to the original contributions [56,111,112]. Certainly, all considered KPIs yield an in-depth overview of the ANN's fitting and predictive capabilities, shedding light on areas of potential improvement, from general trends to specific deviations.

$$R^2 = 1 - \frac{\sum\left(y_{act,i} - y_{pred,i}\right)^2}{\sum\left(y_{act,i} - \overline{y_{act,i}}\right)^2} \quad (6)$$

$$RMSE = \sqrt{\sum\frac{1}{N}\cdot\left(y_{act,i} - y_{pred,i}\right)^2} \quad (7)$$

$$AARD = \frac{1}{N}\cdot\sum\left|\frac{\left(y_{act,i} - y_{pred,i}\right)}{y_{act,i}}\right| \quad (8)$$

$$SD_{av} = \frac{1}{N}\cdot\sum\left(y_{pred,i} - \overline{y_{pred,i}}\right)^2 \quad (9)$$

$$NMAD = \sum\omega_i\cdot\left(\frac{1}{N}\cdot\sum\left|y_{act,i} - y_{pred,i}\right|\right) = \sum\frac{var_{y_i}^{-1}}{\sum_{k=1}^{2}var_{y_k}^{-1}}\cdot\left(\frac{1}{N}\cdot\sum\left|y_{act,i} - y_{pred,i}\right|\right) \quad (10)$$

3. Results and discussion

3.1. Polar soft-SAFT molecular models for new refrigerants

To model the phase equilibria of CO_2-based refrigerant blends using the polar soft-SAFT, molecular models of pure refrigerants are required, inclusive of 38 pure compounds spanning a variety of molecular families. Among these refrigerants, several molecular models have already been developed by our research group, including eighteen refrigerants taken from our previous study [49], in addition to one ether-based single-component, R7200 [113], and a collection of hydrocarbons (R600, R600a, R601, 1-butene) [65,111,112].

This prior work has been completed with new molecular models developed in this contribution for new fluorine-based substances, refrigerants integrating non-standard elements into their carbon-based structures, such as R13I1, effectively incorporating iodine, and RE170, a short-chain ether, first reported in the literature post-2020. Concerning HCs, we have undertaken an accurate recalibration of the ethane-propane and ethylene-propylene series parameters, employing stringent criteria to align with the temperature profiles of air-conditioning systems. This ensures enhanced robustness and applicability within the medium-low temperature ranges, necessary for the advancement of refrigeration and climate control technologies. The re-parameterization

has been carried out by keeping the σ parameter from the works of Vega and collaborators [69,74], and introducing slight refinement in the chain length and dispersive energy values, achieving notable enhancements in the accuracy of computing vapor pressure (with AARD improvements between 86.5 % and 93.7 %). These modifications will also ensure an accurate description of azeotropic behaviors, a critical consideration for subsequent chapters. Presented in Table 3 is the complete list of refrigerant molecules used in this work, offering a concise overview of soft-SAFT parameters, their operational ranges, as well as safety, environmental and technical information.

Although soft-SAFT is a coarse-grained model, its molecular parameters (refer to Table 3) provide information on key structural properties of the molecules that will have an influence on their macroscopic behavior. For a better appreciation, Fig. 2 highlights the combined effects of molecular volume ($m\sigma^3$) and dipole moment (mx_p) contributions for all compounds. For instance, within the di-carbon HFOs, R1123 is distinguished with the largest molecular volume ($m\sigma^3$) and dipole effect, while R1132(E) exhibits greater chain length (m) indicative of its linear conformation. Conversely, R1132a manifests a reduced dipole effect, a consequence of the alignment of fluorine atoms on the same carbon, which reduces the intensity of the molecular polarity in comparison to R1132(E). R1234ze(E) stands out among the tri-carbon HFO family members with its long chain length, a consequence of a more linear structural conformation, while R1234ze(Z) is notable for a bulkier size, evidenced by the largest segment diameter and molecular volume. R1234yf and R1243zf present an intermediate pattern, balancing molecular dimensions and reduced polarity due to a more localized effect of the dipole.

R1233zd(E) has longer chain lengths, contrary to R1224yd(E), which exhibits a larger segment diameter, influenced by the spatial arrangement of the chlorine amidst the fluorine atoms. The dipole segment value for R1224yd(E) exceeds that of R1233zd(E), reflecting the complete bonding of its carbon structure to electronegative elements, either strongly polarizing fluorines or moderately electronegative chlorines. For carbon-based HFOs with four carbons, R1336mzz(E) shows the greatest chain length (m), contrasted by R1336mzz(Z)'s higher volume ($m\sigma^3$) and segment diameter (σ), placing R1336yf and R1345fz in a median range. The chain length of R1336mzz(E) was adjusted to exceed that of R1225ye(Z), R1233zd(E), and all stereoisomers of R1224yd, ensuring meaningful physical values. The x_p value, indicative of the dipole's influence, remains constant across all isomers, fixed to maintain uniformity in the dipole's influence across different molecular configurations. Regarding the dipolar contributions of RE170 and R13I1, particular attention is given to the electronegativity of constituent atoms, which plays a key role in determining the strength of the dipole effects. The hierarchy of electronegativity (fluorine > oxygen > chlorine) guides the impact of dipolar interactions, as iodine and hydrogen's minimal contributions have led to a dipole-affected fraction of 3/5 for R13I1, reflecting the iodine atom's dimensions within the molecular volume. For RE170, the relatively similar atomic weights of oxygen and carbon, unlike the heavier halogen atoms, justify a minimal ratio of 1/3. Further validation involved delving into alkanes and ethers, such as

Table 3
Polar soft-SAFT molecular parameters for refrigerants assessed in this work in addition to Absolute Average Deviations (AAD%) for vapor pressure (P) and saturated liquid density (ρ).

Name	ASHRAE (GWP_{AR5})	m	σ [Å]	ε [K]	μ^* [C·m]	x_p	T [K]	AAD_P [%]	AAD_ρ [%]	Ref
R41	A2 (116)	1.371	3.400	180.3	6.1743	0.50	200–280	1.494	0.763	[49]
R32	A2L (677)	1.376	3.506	164.5	6.5979	0.75	200–316	0.565	0.255	[49]
R23	A1 (12400)	1.397	3.610	147.9	5.5005	0.90	200–270	1.327	0.155	[49]
R161	A2 (4.0)	1.577	3.693	232.3	6.4701	0.33	200–340	1.374	0.414	[49]
R152a	A2 (138)	1.662	3.754	202.3	7.5452	0.50	200–356	0.993	0.792	[49]
R134a	A1 (1300)	1.813	3.770	169.5	6.8648	0.70	200–344	1.443	0.389	[49]
R125	A1 (3170)	1.887	3.790	165.1	5.2136	0.90	200–310	1.618	0.274	[49]
R245fa	B1 (858)	2.479	3.675	197.1	5.1669	0.80	230–395	1.593	0.596	[49]
R236fa	A1 (8060)	2.056	4.012	172.4	6.6112	0.90	270–370	2.410	0.554	[49]
R227ea	A1 (3350)	2.131	4.033	190.7	4.8567	1.00	230–345	4.224	0.260	[49]
R1123	A2L (3.0)	1.527	3.760	175.3	5.7373	0.80	230–300	2.963	0.199	[49]
R1132a	A2 (<1.0)	1.750	3.530	160.9	4.5998 [113,114]	0.65	198–293	1.676	0.222	This Work
R1132(E)	NA (1.9)	2.598	2.988	179.9	-	-	220–325	1.289	0.309	This Work
R1225ye(Z)	A/B1 (<1.0)	2.077	3.845	172.4	6.0375	0.8	250–355	1.018	0.260	[49]
R1234yf	A2L (<1.0)	1.740	4.082	191.6	6.7079	0.70	250–331	1.150	0.314	[49]
R1234ze(E)	A2L (1.0)	2.044	3.821	204.0	4.8033	0.75	250–351	3.016	0.555	[49]
R1234ze(Z)	A2L (1.4)	1.541	4.271	180.7	9.6734 [34]	0.75	240–360	1.735	0.272	This Work
R1243zf	A2 (1.0)	1.904	3.880	170.0	8.1689	0.50	270–345	1.092	0.295	[49]
R1336mzz(E)	A1 (14.0)	3.633	3.386	184.6	-	-	286–375	1.847	0.605	This Work
R1336mzz(Z)	A1 (2.0)	1.806	4.430	195.6	10.641	0.60	325–415	3.287	0.565	[49]
R1336yf	NA (NA)	2.047	4.180	155.7	8.4175 [115]	0.70	278–353	2.112	0.546	This Work
R1345fz	NA (NA)	2.031	4.150	143.6	9.4282 [115]	0.60	278–353	1.705	0.643	This Work
R1233zd(E)	A1 (5.0)	2.331	3.819	232.6	3.8119	0.80	250–400	2.936	0.506	[49]
R1224yd(E)	NA (NA)	2.171	3.957	190.5	6.8614 [115]	0.85	300–380	2.533	0.440	This Work
R1224yd(Z)	A1 (1.0)	2.278	3.899	202.4	5.6372	0.85	280–375	1.358	0.237	[49]
R170	A3 (6.0)	1.413	3.756	204.9	-	-	150–280	2.136	0.822	This Work
R290	A3 (3.3)	1.801	3.811	222.4	-	-	150–340	1.470	0.396	This Work
R600	A3 (4.0)	2.134	3.871	237.7	-	-	NA	NA	NA	[69]
R600a	A3 (3.0)	1.942	4.036	240.3	-	-	200–400	NA	NA	[116]
R601	A3 (11.0)	2.497	3.901	246.6	-	-	NA	NA	NA	[69]
R601a	A3 (11.0)	2.223	4.068	254.9	-	-	180–380	2.465	0.318	This Work
R1150	A3 (4.0)	1.711	3.372	170.5	-	-	180–270	1.889	0.446	This Work
R1270	A3 (2.0)	1.979	3.566	207.9	-	-	180–340	1.830	0.264	This Work
1-butene	NA (NA)	2.147	3.762	228.8	-	-	120–420	NA	NA	[117]
CO_2^+	A1 (1.0)	1.571	3.166	166.5	-	0.33	-	0.340	0.562	[44]
R13I1	A1 (1.0)	2.004	3.830	223.0	3.4958 [34]	0.60	300–360	0.220	0.287	This Work
RE170	A3 (1.0)	2.009	3.447	225.0	4.3397 [34]	0.33	180–350	1.099	0.168	This Work
R7200	NA (57.0)	2.113	4.916	214.6	11.2100	0.68	280–380	3.243	0.162	[118]

* ·10E-30
+ 14.68E-40C·m^2

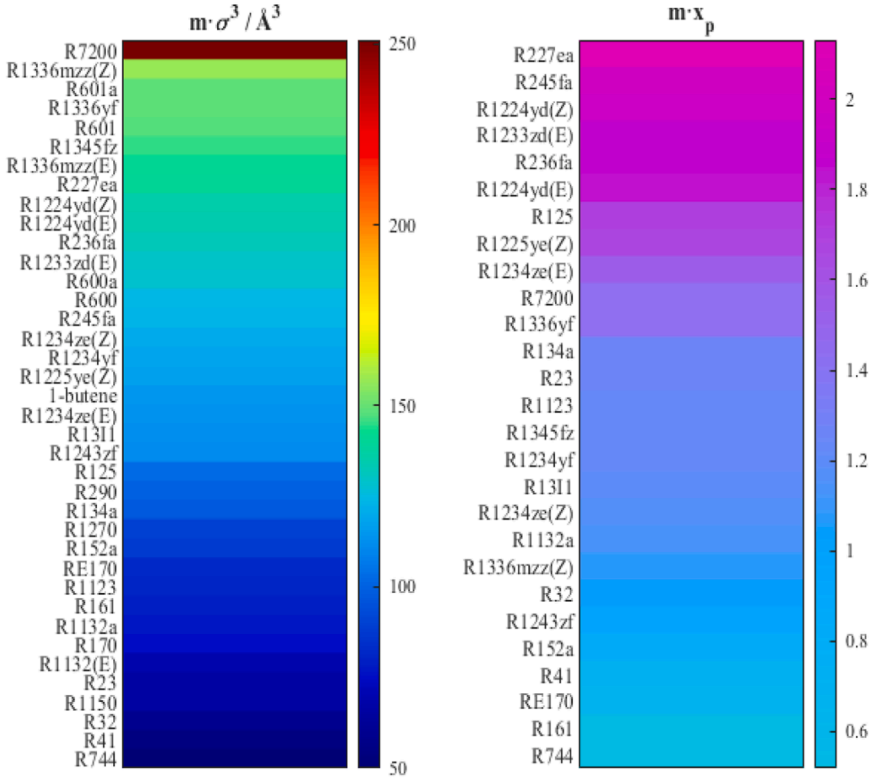

Fig. 2. Heatmap of the combined effects of molecular volume ($m\sigma^3$) and dipole moment (mx_p) contributions for all compounds modeled in Table 3.

R290 and diethylene glycol dimethyl ether, respectively, establishing RE170's chain length at a midpoint.

The dipole moment contribution (mx_p) in Fig. 3**a** correlates with molecular weight across 20+ refrigerants, spanning dipole and quadrupole types. Alongside, the recalibration's precision within the hydrocarbon series is highlighted by the molecular volume's robust correlation with molecular weight, demonstrated by an R^2 of 0.997 for saturated hydrocarbons, and surpassing 0.999 for the hydroolefin-based compounds under assessment, as appreciated in Fig. 3**b**. Probing further into the molecular complexities, the chain length and segment diameter reveal an ordered progression consistent with molecular size: R601 exhibits the longest chain, followed sequentially by R601a, R600, R600a, R290, and R170. In terms of segment diameter, a similar order is apparent, with R601a displaying the largest value, moving down through R600a [116], R601, R600, R290, to R170 [69]. Significantly, 1-butene is characterized by larger chain lengths and segment diameters compared to its olefin-based counterparts.

Provided in Fig. 4 is a comparison of the vapor pressure, saturated densities and latent heat for selected novel pure refrigerants computed using polar soft-SAFT EoS and experimental data (see Table S1 in the Supplementary Material for details) [34,115,119–125]. The excellent quantitative agreement between model calculations and experimental data, maintaining an AAD% below 3 % for vapor pressure and under 1 % for saturated densities, reinforces the molecular models' robustness and accurate representation of these fluids. The VLE diagram includes predictions of the critical point, overestimated in this case due to excluding

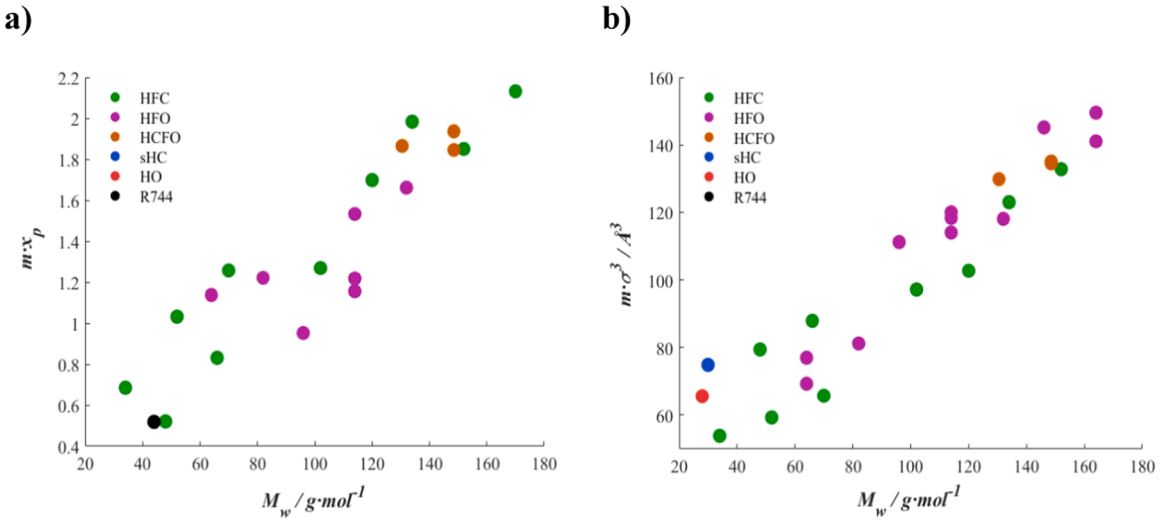

Fig. 3. Correlation of molecular weight (M_w) on **a)** dipole moment, and **b)** molecular volume for selected refrigerants in Table 3.

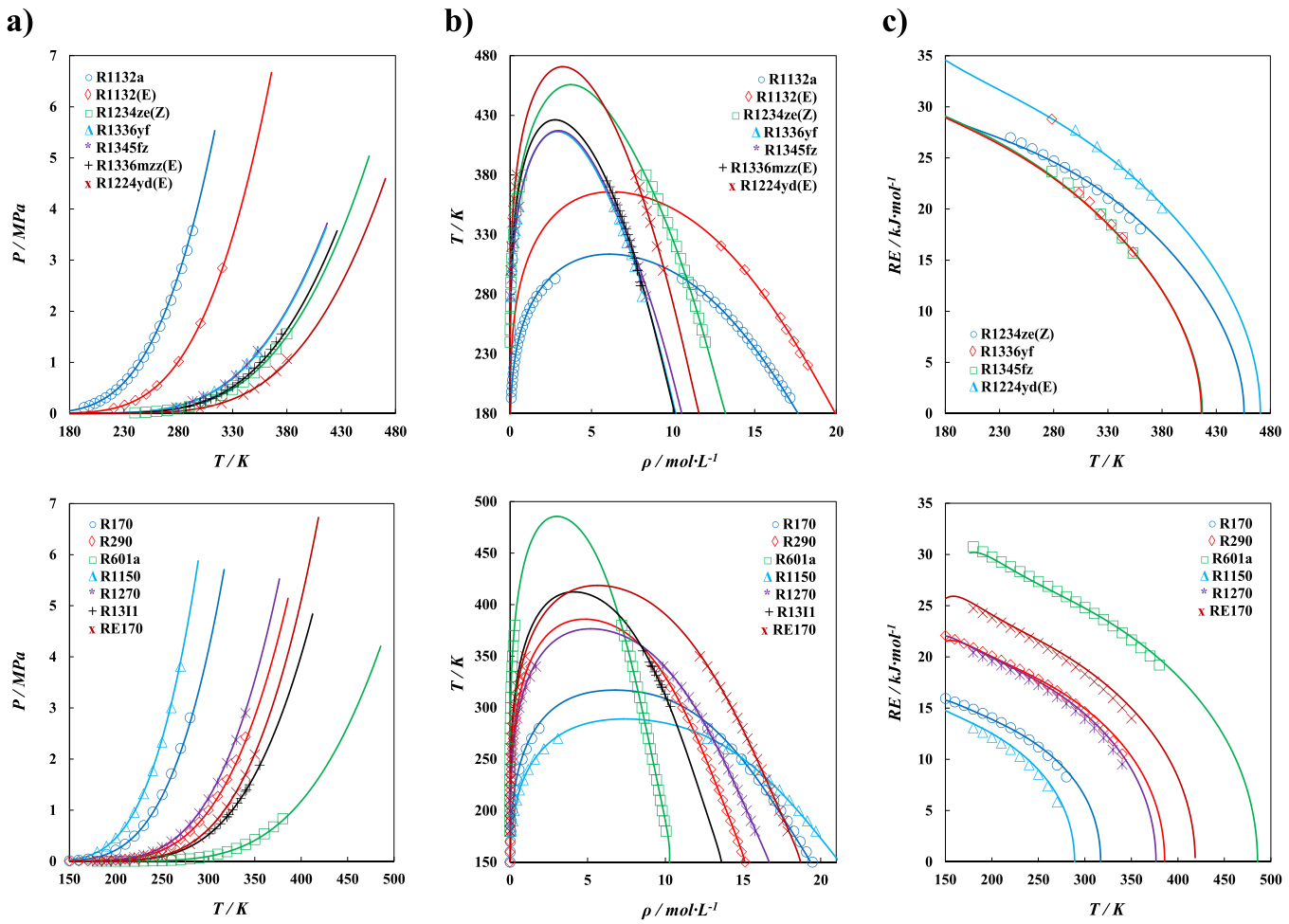

Fig. 4. **a)** Vapor Pressure, **b)** Saturated Density, and **c)** Refrigeration effect (latent heat) diagrams of novel pure refrigerants listed in Table 3. Symbols correspond to experimental data (see references in the text) and lines to polar soft-SAFT calculations.

effect of long-range density fluctuations, corrected with the equation's crossover term [126]. This is omitted, as the selected temperature range meets sub-critical air conditioning requirements for CO_2-based mixtures in cascade cycles. In Fig. 4**c**, the refrigeration effect or latent heat is accurately estimated within an AARD of 3 %, except for R1150 and R1270, whose AARDs are slightly higher (4.84 % and 3.18 %, respectively). In particular, R601a and R1224yd(E) stand out for their exceptional accuracy, with deviations below 1 %, closely followed by R1234ze(Z) and R1336yf. However, it must be noted that R1336yf's outlier data at 280 K has been omitted in the AAD calculation due to its inconsistency with the generic trend.

Fig. 5 further confirms the model's reliability, predicting the binary blend of R13I1 and R152a with excellent agreement across tested isotherms, precisely capturing the azeotropic behavior in alignment with experimental evidence. The observed azeotropic behavior is attributed to the similarity in normal boiling points of both single components, with values of 250.65 K for R13I1 and 249.15 K for R152a, along with comparable fractions of the molecules affected by the dipole moment—0.6 for R13I1 and 0.5 for R152a—which enhance the interactions between the two components. Therefore, the coarse-grain model developed for R13I1 in this work, in conjunction with the model for R152a derived from previous work [49], yields quantitative predictive performance for the mixture VLE without the need to fit any binary interaction parameter.

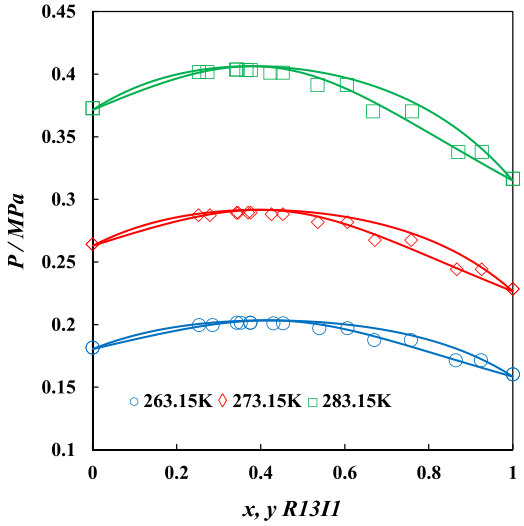

Fig. 5. VLE of R13I1 + R152a binary mixture at 263.15 K (blue circles), 273.15 K (red diamonds), and 283.15 K (green squares). Symbols correspond to experimental data [127] and lines to polar soft-SAFT characterizations.

3.2. Characterization of CO_2-based refrigerant blends using polar soft-SAFT

The study of binary systems is completed by a comprehensive screening of CO_2-based mixtures. Of the 38 pure-components displayed in Table 3, 27 CO_2-based systems with existing experimental data in the literature are considered in Fig. 6 and Figure S1 in the Supplementary Material. These Figures display the VLE curves of CO_2 + refrigerant binary blends computed using polar soft-SAFT EoS. In few cases, these binary mixtures are predicted directly from the model setting binary parameters to unity, while in other cases, a constant temperature-independent binary energy parameter (ξ_{ij}) is regressed to a single isotherm for each studied mixture, as specified in Table 4. In all cases, good agreement between the polar soft-SAFT calculations and experimental data is noted, appropriately accounting for variations in the geometric average dispersive energy. Further examination of Table 4 reveals that, while a third of combinations are predictively described (i.e. $\xi_{ij} = 1.00$), a majority of blends require minor adjustments from ideality ($0.98 < \xi_{ij} < 1.02$) for precise modeling. This suggests the model's robustness in fine-tuning the solubility of CO_2 in selected refrigerants of different nature.

The VLE results indicate azeotropic behavior when CO_2 is combined with R170 or R1150 (see Fig. 6), corroborated by empirical evidence across various isotherms. A positive deviation from Raoult's law is particularly visible in the solubility behavior of CO_2 with HFCs, HFOs, HCFOs, HCs, and RE170, where strong attractive interactions are taking effect. The analysis of P_{xy} solubility diagrams indicates stronger interactions of CO_2 with R32, R1336mzz(E), and RE170 (see Figure S1) than with other HFC and HFO refrigerants, a phenomenon linked to the increased asymmetry in its molecular structures. In contrast, for saturated hydrocarbons R13I1 and R7200, the trend inverts, indicating more repulsive forces than the simple geometric average within the mixture components, suggestive of van der Waals forces playing a predominant role in the phase behavior. This dichotomy illustrates the intricate interplay of molecular forces and the capability of the SAFT model in accounting for these behaviors.

3.3. Binary mixtures data development through artificial neural networks

3.3.1. Design of ANN configuration with optimal number of neurons

The proposed AI algorithm [128,129] is based on the charge distribution molecular descriptors obtained from COSMO-RS. Consequently, an in-depth evaluation of the electrostatic potential of molecular surfaces according to σ-profiles for various refrigerants and blends was carried out. In this context of data assembly and preprocessing, a wide-ranging array of data was compiled, serving as outputs to the AI-model, consisting of 531 data points for vapor-liquid equilibria of CO_2-based refrigerant blends, where 423 were used for training (80 %), 54 for validation (10 %), and 54 for testing (10 %). Specifically, this collection features 494 data points covering CO_2 compositions (both in liquid and vapor phases) ranging from 5 % to 95 % in mole fraction. However, R7200 is excluded from consideration, due to extreme saturation pressure divergence from the comparative baseline established by commercial air conditioning standards. Additionally, the pronounced temperature glide with CO_2 [37]—irrespective of pressure and composition—renders this particular mixture less relevant, especially since it lies outside the targeted medium-low temperature range for air conditioning applications central to this contribution. The output database is

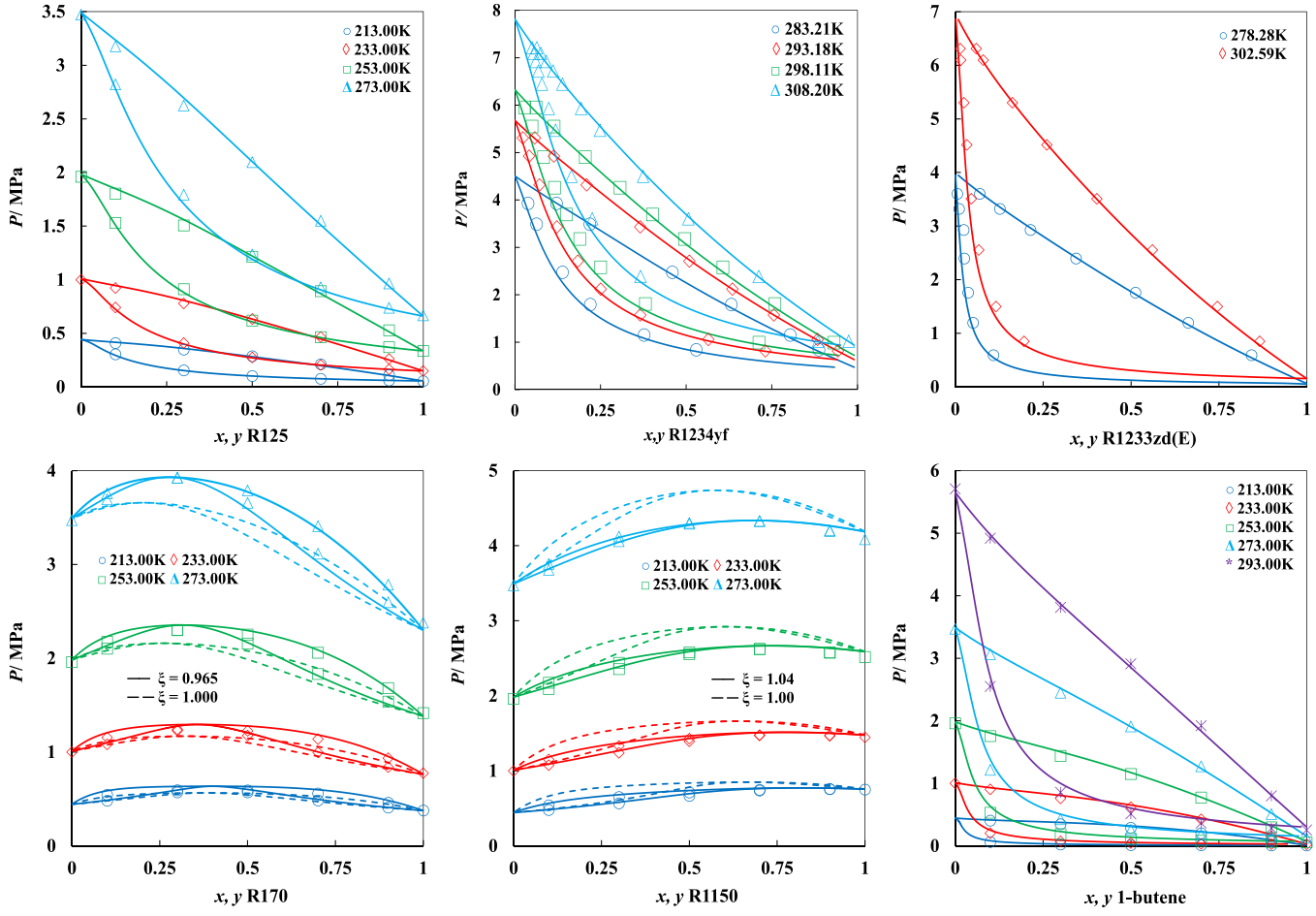

Fig. 6. Selected VLEs for CO_2-based binary mixtures. Symbols correspond to experimental data and solid lines to polar soft-SAFT characterizations.

Table 4
Temperature-independent binary parameter, ξ_{ij}, employed for each CO_2-refrigerant pair, regressed to a single VLE isotherm for each mixture. A value of $\xi_{ij} = 1.00$ means that no corrections are necessary.

CO_2+	ξ_{ij}	T/K	Ref.	CO_2+	ξ_{ij}	T/K	Ref.	CO_2+	ξ_{ij}	T/K	Ref.
R41	1.02	240	[34]	R1243zf	1.00	289	[36]	R600	0.96	253	[34]
R32	1.05	244	[34]	R1234yf	1.00	293	[91]	R600a	0.96	253	[34]
R23	1.03	263	[34]	R1234ze(E)	1.00	293	[92]	R601	0.98	253	[34]
R161	1.03	263	[34]	R1234ze(Z)	1.02	293	[38]	R601a	0.97	253	[34]
R152a	1.00	278	[34]	R1336mzz(E)	1.05	293	[38]	R1150	1.04	253	[34]
R134a	1.03	273	[34]	R1233zd(E)	1.00	278	[39]	R1270	1.01	253	[34]
R125	1.00	253	[34]	R7200	0.97	313	[37]	1-butene	1.00	253	[34]
R227ea	1.00	273	[34]	R170	0.97	253	[34]	R13I1	0.98	253	[93]
R1123	1.02	270	[35]	R290	0.97	253	[34]	RE170	1.05	253	[34]

completed with the saturation targets for all 27-constituent pure-components (including pure CO_2), supplemented by 10 additional F-based single refrigerants. These include R245fa, R236fa, R1225ye(Z), R1336mzz(Z), R1132(E), R1132a, R1345fz, R1336yf, R1224yd(E), and R1224yd(Z). For a more complete understanding of the σ-profiles of selected refrigerants, the reader is referred to previous work [60].

The effectiveness of ANN for predicting bubble and dew point temperatures was assessed with two hidden layers, each containing 1 to 25 neurons, treating the number of neurons in each hidden layer as an adjustable parameter. The analysis, based on the RMSE indicated that optimal network configurations were associated with higher neuron counts in both hidden layers, lowering weighted RMSEs, especially for dew temperature predictions. In contrast, configurations with a lower degree of complexity consistently yielded RMSEs surpassing a threshold of 5 units when hyper-tuning the neuron counts in both hidden layers. Consequently, the optimal artificial neural network configurations for predicting bubble (24 HL1 – 20 HL2) and dew temperatures (23 HL1 – 15 HL2) were identified through empirical analysis, with respective weighted RMSEs of 0.0388 and 0.2003. Such findings, after an extensive fine-tuning process (see Fig. 7), emphasize the necessity of a larger architecture in both hidden layers to achieve precision in ANN-based predictions, with an average elapsed time per epoch of 13 seconds. Cross-validation was evaluated by testing 625 ANN configurations per output, with five runs per configuration using random splits for training, validation, and testing. In this manner, the best-performing output was consistently achieved across all runs and on average, highlighting the model's stability and robustness.

Fig. 8 depicts the integrated multitask ANN architecture, optimized for predicting saturation temperatures at 1 MPa, which emerged as the most suitable model after extensive testing and adjustments. This selected ANN layout, incorporating a multi-output framework, comprises a total of 3727 parameters and 84 biases, with 1964 and 1763

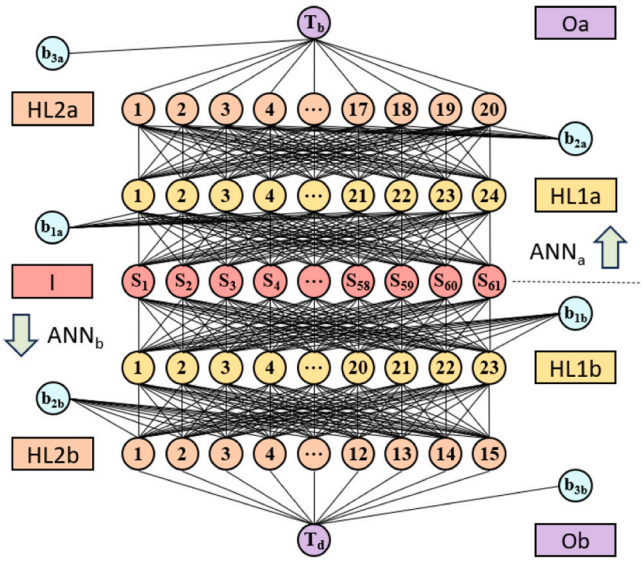

Fig. 8. Best-performing multitask model for predicting bubble and dew temperatures of CO_2-based refrigerant mixtures, showcasing weights, and biases interconnections within the integrated ANN architecture.

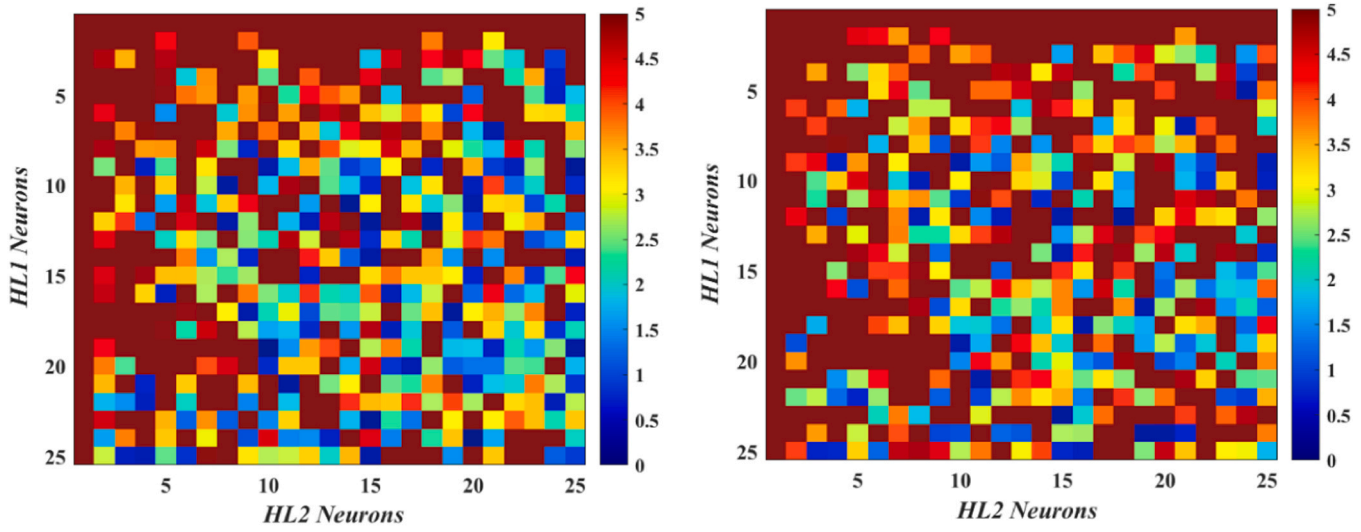

Fig. 7. RMSE colorbar from 0 (intense blue) to 5 (intense red) in relation to neuron counts in hidden layers 1 (HL1) and 2 (HL2), ranging from 1 to 25, for the calculation of the bubble (left) and dew (right) temperatures.

weights, and 45 and 39 biases specifically allocated for bubble and dew temperatures layouts, respectively.

The ensemble ANN setups are outlined in Tables S2–3, where corresponding weights, biases, and output parameters streamline the fine-tuning of saturation properties for a range from pure to binary CO_2-based refrigerant blends. Eq. (11) depicts the bubble point output function, while the dew saturation temperature can be estimated by Eq. (12). It is important to ensure the relevance of maintaining CO_2 within its triple (216.55 K, 0.5185 MPa) and critical points (304.13 K, 7.377 MPa) for effective refrigeration usage, ensuring ease of liquid-vapor transition (appreciate this phenomenon in Fig. 6 and **S1** in the Supplementary Material). Hence, to prevent CO_2-based mixtures from solid state occurrence, the working pressure is standardized at 1 MPa, aligning with CO_2's unique phase properties and ensuring its viability in sub-critical cascade refrigeration technologies.

$$\begin{aligned} T_b/K = {} & 1.0585(HH_1) - 0.2395(HH_2) - 1.3856(HH_3) - 0.5196(HH_4) \\ & + 0.5345(HH_5) + 0.8107(HH_6) + 0.8575(HH_7) + 1.0926(HH_8) \\ & - 0.2386(HH_9) - 1.8389(HH_{10}) - 0.4634(HH_{11}) \\ & - 0.7455(HH_{12}) + 0.5813(HH_{13}) + 0.5040(HH_{14}) \\ & + 0.6480(HH_{15}) + 1.9120(HH_{16}) + 0.6951(HH_{17}) \\ & - 0.5351(HH_{18}) - 0.7133(HH_{19}) - 0.3663(HH_{20}) + 0.9781 \end{aligned} \tag{11}$$

$$\begin{aligned} T_d/K = {} & -1.6300(HH_1) - 0.9403(HH_2) + 0.4453(HH_3) - 0.5254(HH_4) \\ & + 1.8056(HH_5) + 1.3672(HH_6) - 1.4100(HH_7) + 1.2360(HH_8) \\ & - 4.0551(HH_9) - 0.6360(HH_{10}) - 1.3384(HH_{11}) \\ & - 3.1070(HH_{12}) - 1.2178(HH_{13}) - 0.5708(HH_{14}) \\ & + 1.5406(HH_{15}) + 0.0370 \end{aligned} \tag{12}$$

3.3.2. Optimal run evaluation

Based on the previous results for optimizing the performance of the ANN architect (see Fig. 7), the optimal ensemble ANN setup is rigorously evaluated using a blend of visual tools, such as parity plots and distribution patterns in Fig. 9**a**, standardized residuals in Fig. 9**b**, and statistical methods detailed in Table 5.

The parity plots in Fig. 9**a** depict a graphical representation of the model's performance, linking the relationship between the ANN-predicted saturation temperatures and the soft-SAFT data used as training outputs. For the liquid phase transition threshold, the majority of training data lies within a narrower temperature range of 200 – 300 K, accounting for over 90 % of the validation and testing datasets. The model's high accuracy is a direct consequence of the distribution pattern within this range, as seen in the high degree of alignment along the $y = x$ diagonal. However, the developed ANN exhibits a reduced accuracy for predictions above 350 K, as evidenced by the sole data point in this range showing significant deviation. In contrast, dew point predictions,

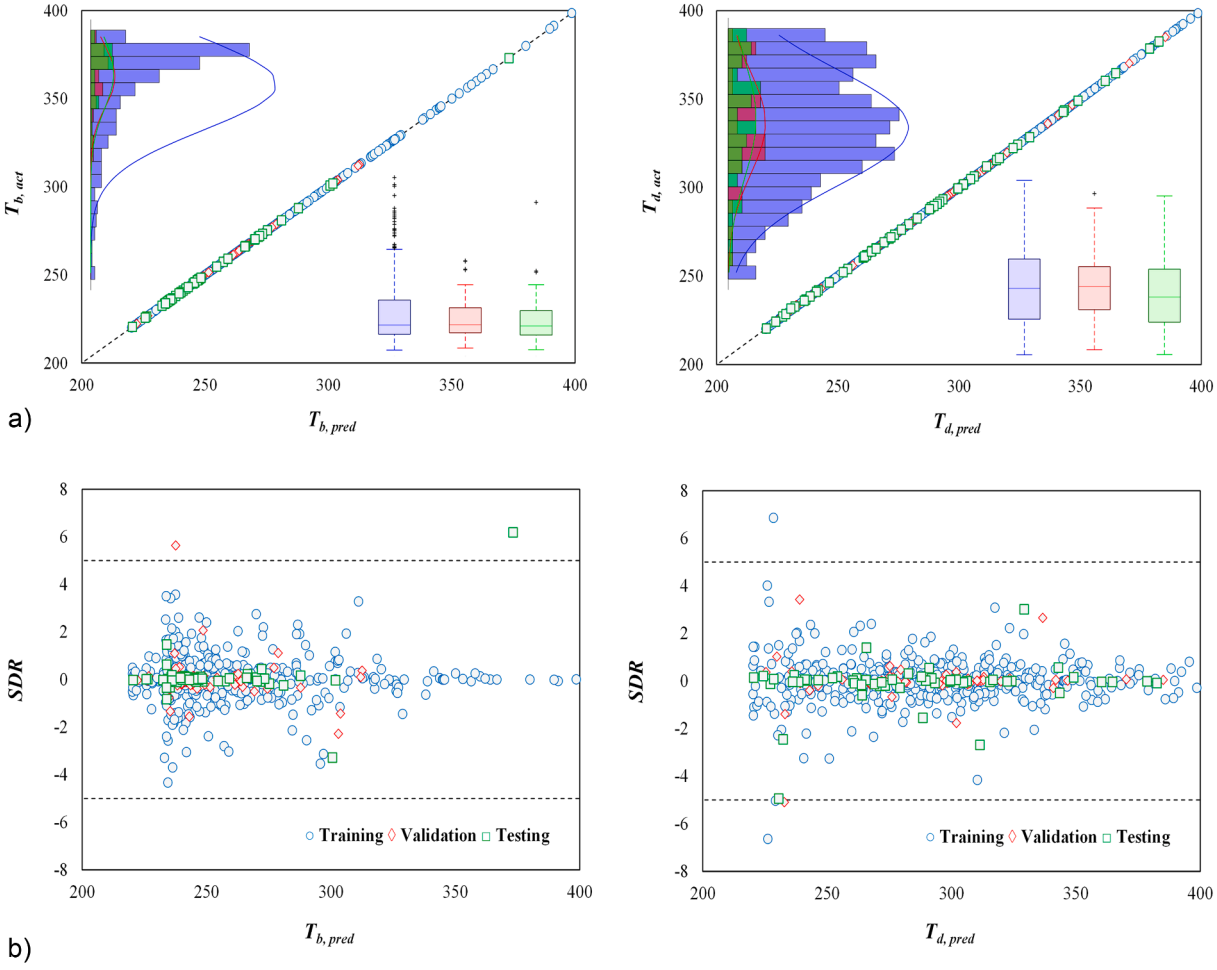

Fig. 9. a) Scatter, distribution patterns, box, and **b)** Standardized residual diagrams (SDR) for training (blue), validation (red), and testing (green) datasets and in correlation with both observed and predicted responses at 1 MPa. The distribution plot displays the probability distribution of data in histograms, illustrating how values are spread across the datasets, while the box plot provides insights into central tendency and variability by depicting the median, quartiles, and outliers, ranging from minimum to maximum data points.

Table 5
Statistical KPIs for the developed ANN models in predicting dew and bubble points.

Output	Metric	Training	Validation	Testing	Total
Bubble	R^2	1.0000	1.0000	1.0000	1.0000
Dew		1.0000	1.0000	1.0000	1.0000
Bubble	RMSE	0.0016	0.0135	0.0414	0.0172
Dew		0.0213	0.1691	0.2291	0.0959
Bubble	AARD/%	0.0004	0.0026	0.0045	0.0024
Dew		0.0049	0.0270	0.0356	0.0220
Bubble	SD_{av}	0.00077	0.0046	0.0092	0.0020
Dew		0.00953	0.0498	0.0680	0.0195
Both	NMAD	0.00018	0.0012	0.0018	0.0004
Both	Data Points	**423**	**54**	**54**	**531**

trained on a more diverse temperature range including over 30 % of data below 250 K and about 9.4 % above 350 K, display higher overall residuals indicative of a slightly lower extrapolative capacity. This is particularly noticeable at temperatures above 300 K, where the model has been trained with a broader spectrum of data, a segment constituting 40.7 % of the training data, as depicted in Fig. 9**b**. Partly because of this basis, and further influenced by the substantially higher standard deviation of the vapor phase database at equilibrium (see Fig. 9**b**), the overall residuals across all datasets (*i.e.*, training, validation, and testing) show an increase relative to the bubble point. Enhancing the ANN's complexity or expanding the dataset could further improve the model's projections on dew temperature outcomes, albeit challenged by the limited availability of CO_2-based blend data in literature.

Statistical analysis, shown in Table 5, provides a quantitative backbone to these observations. The KPIs such as R^2 values, exceeding 0.9999 for all cases, in addition to NMAD, indicate the robustness of the developed neural network. However, a closer look at other KPIs reveals subtle differences in performance, especially in scenarios beyond the 300 K range. While bubble point predictions maintain high precision across datasets, the estimation of the dew phase exhibits slightly lower performance metrics, particularly when moving from training to validation and testing stages. These observations are directly related to data diversity and distribution in training ANN models. The model's accuracy in bubble points, trained on a 200 – 300 K range, contrasts with the broader range of temperatures in the vapor phase, demanding a more elaborated architecture to handle the increased complexity. Otherwise, this analysis still shows the efficacy of the ANN model in predicting saturation temperatures for dew and bubble points, even though with a notable distinction in performance due to data diversity.

3.3.3. Testing the validity of ANN-based model for predicting saturation properties of CO_2-based refrigerant blends

Fig. 10 serves as a robust validation to the multitask ANN's capability in accurately modeling the saturation properties of diverse CO_2-based refrigerant blends. Employing a fine-grained discretization approach with a 0.1 % composition interval, we have computed each mixture's corresponding COSMO descriptor as inputs to the ANN to describe the VLE. A significant feature is the accurate description of azeotropic behaviors in CO_2 mixtures involving short-chained hydrocarbons (R170 and R1150), in contrast with the zeotropic nature observed in mixtures containing longer, bulkier compounds. Validation and testing data closely match polar soft-SAFT output data used in ANN development, with the exception in the dew phase of the CO_2 + R1150 mixture at low fractions of the light hydrocarbon, where unphysical behavior is obtained. This deviation – characterized as an underestimation – was anticipated from the residual plots in Fig. 9**b**, where the testing point on the edge of the SDR limit of –5 falls outside the scope of the ANN's optimal range, indicating a limitation in its applicability. Special focus is needed on the dew phase, where three training points have SDRs

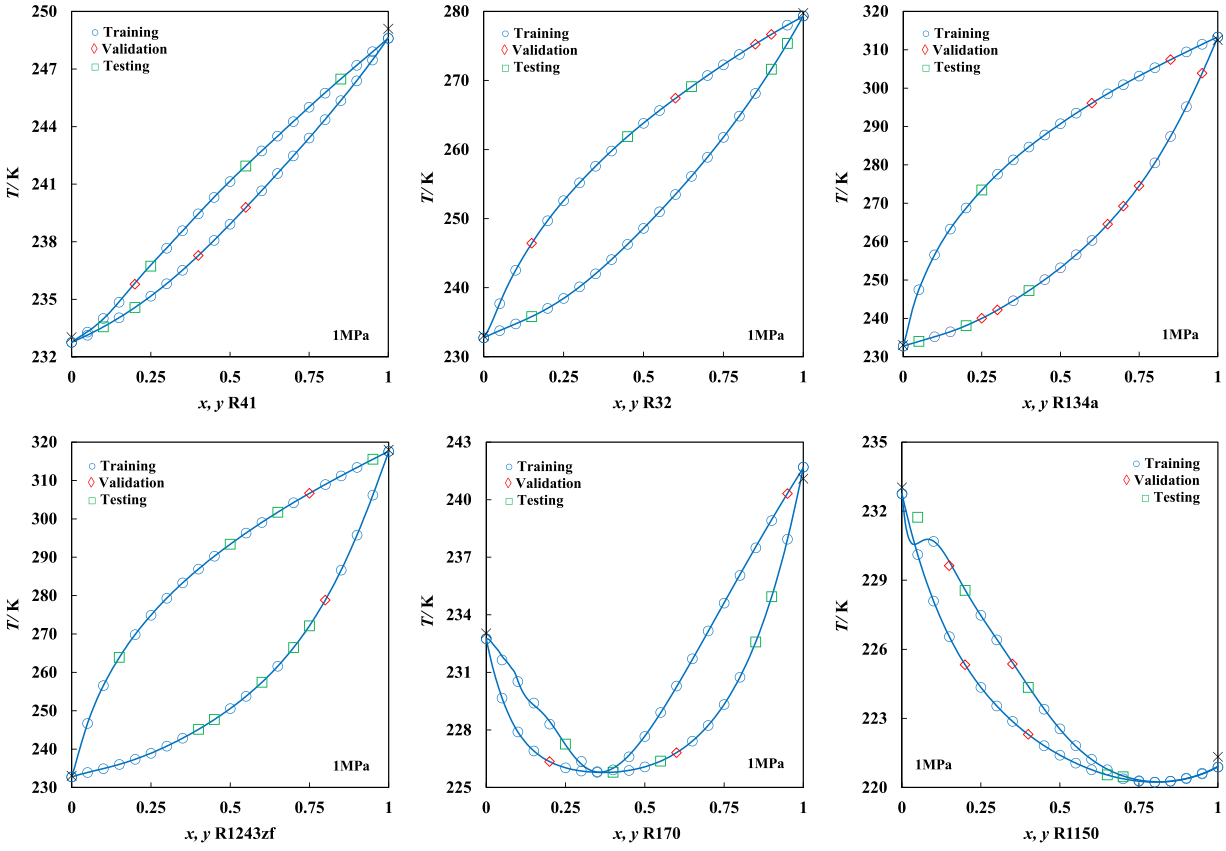

Fig. 10. Temperature-composition diagrams of selected CO_2-based blends at 1 MPa. Symbols correspond to polar soft-SAFT data used for training (blue circle), validation (red diamond), or testing (green square), while lines stand for ANN calculations.

exceeding ± 5 and twelve points surpass an SDR of 3, indicating the need for prudent interpretation around such extreme ranges, particularly for the case of an unphysical output. The discrepancy in the training point for the dew phase of the R170 and CO_2 mixture at 15 % mole composition, with an SDR near 3, exemplifies the minor, yet visible deviation observed. While outliers are present, the ANN's overall fitting closely matches SAFT outcomes across diverse mixtures, demonstrating its reliability through early deviation detection. Experimental saturation temperatures for pure substances from NIST, plotted with polar soft-SAFT estimates, are in excellent agreement with the ANN's outputs in all cases.

3.4. Practical application of the ANN developed model for fine-tuning polar soft-SAFT binary parameters

Based on the good descriptive nature of the presented ANN, the saturation properties of new CO_2-based blends, including HFOs and HCFOs without existing experimental or molecular simulation data, have been predicted. Given previous results, the approach enables the fitting of polar soft-SAFT binary parameters without relying on experimental data. These findings, presented in Fig. 11, reveal the ANN's capability in forecasting saturation profiles for both liquid and vapor phases, leading to precise determination of binary parameters for selected F-based compounds, with R1336mzz(Z) and R1224yd(E) yielding ξ_{ij} values of 1.02 and 1.03, respectively. Such domains align with established literature on CO_2 mixtures (see Table 4), consistently including HFCs, HFOs, and HCFOs.

The characterization is validated by comparison with the NRTL activity coefficients model, and demonstrates the consistency of the developed methodology, as evidenced by the AARD obtained for both mixtures respect to this model. Specifically, the CO_2 mixture with R1336mzz(Z) (see Fig. 11**a**) is described with an overall AARD close to 3 % for bubble temperatures and below 1 % for dew points, indicating highly accurate predictions. Likewise, the mixture with R1224yd(E) (see Fig. 11**b**) shows deviations below 3.5 % for both measures. In the liquid saturation phase, as detailed in Table 6, the ANN demonstrates enhanced predictive capability at mole concentrations close to pure CO_2, highlighting the need to expand the dataset to include a broader spectrum of 3rd and 4th generation refrigerant compounds and blends. Moreover, it is observed that bubble predictions deviate from actual data within the upper range of molar composition for the F-based compound.

Table 6
Actual (NRTL) and predicted (ANN) bubble and dew temperatures in addition to the Average Absolute Relative Deviation (AARD) for the CO_2-based mixtures depicted in Fig. 11.

x_1, y_1	T_b^{NRTL} / K	T_d^{NRTL} / K	T_b^{ANN} / K	T_d^{ANN} / K	ARD T_b	ARD T_d
			(1) R1336mzz(Z) / CO_2			
0.9	312.63	383.95	348.53	384.92	11.48	0.25
0.8	285.94	378.80	305.41	379.39	6.81	0.15
0.7	270.90	373.17	276.90	373.10	2.21	0.02
0.6	260.86	366.92	263.26	365.99	0.92	0.25
0.5	253.50	359.86	256.36	357.98	1.13	0.52
0.4	247.76	351.63	251.15	349.19	1.37	0.69
0.3	243.11	341.66	246.33	340.68	1.33	0.29
0.2	239.22	328.77	241.87	332.85	1.11	1.24
0.1	235.90	309.41	237.81	317.46	0.81	2.60
			(2) R1224yd(E) / CO_2			
0.9	308.47	372.61	341.54	372.61	10.72	1.65
0.8	284.42	365.53	308.61	365.53	8.50	1.15
0.7	270.20	359.21	284.39	359.21	5.25	0.97
0.6	260.49	353.70	267.89	353.70	2.84	1.22
0.5	253.29	348.12	256.81	348.12	1.39	1.70
0.4	247.64	340.89	249.23	340.89	0.64	2.05
0.3	243.04	332.38	243.73	332.38	0.28	2.54
0.2	239.19	325.79	238.98	325.79	0.08	4.62
0.1	235.88	313.30	233.62	313.30	0.96	7.07
				$AARD_1$ / %	3.02	0.67
				$AARD_2$ / %	3.41	2.55

This discrepancy is attributed to insights from Fig. 9, indicating the inadequacy of the ANN in capturing saturation temperatures within this specific temperature range, primarily due to limitations in the training dataset.

Certainly, effective liquid phase characterization is essential for accurate SAFT descriptions ensuring reliable VLEs, particularly in highly zeotropic mixtures akin to Fig. 11's. This phenomenon, evident in CO_2-based mixtures from long-chain hydrocarbons (R600 to R601a) to F-based coolants like HFOs (R1336mzz(E), R1234ze(Z)), extending also to R13I1, results in a more subtle pressure gradient effect for the vapor phase when applying energy-based binary parameters.

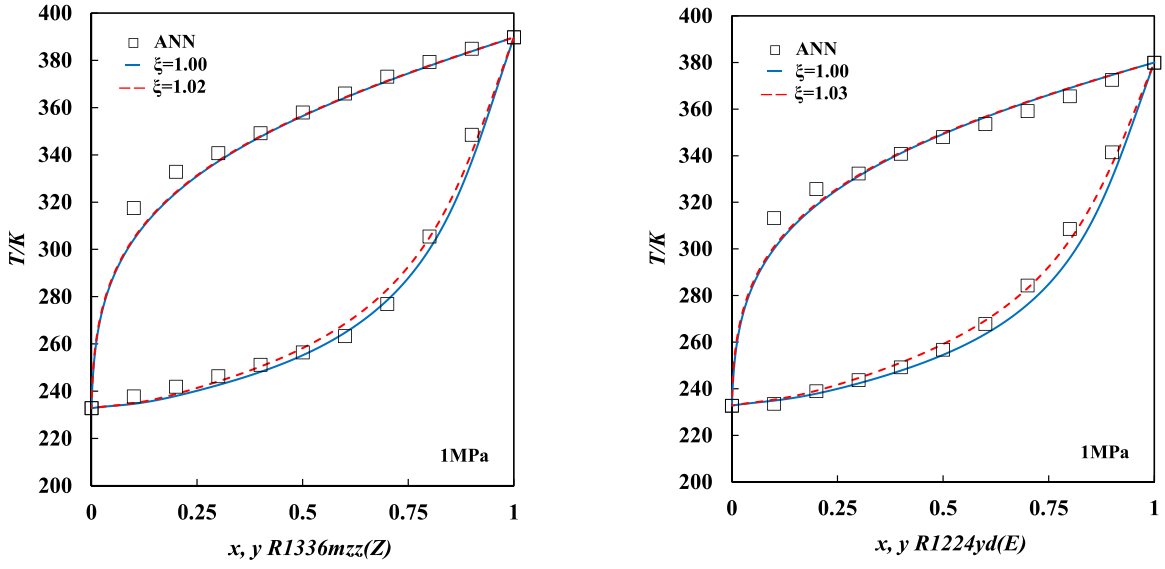

Fig. 11. Temperature-composition diagrams of selected CO_2-based blends excluded from ANN fitting or external validation at 1 MPa. Symbols correspond to saturation data obtained from ANN calculations, while lines are soft-SAFT characterizations. Note that characterizations using an adjustable fitted ξ_{ij} parameter is shown in red discontinuous patterns, while pure SAFT predictions are depicted with blue solid lines.

4. Conclusions

A machine learning-based methodology has been developed to predict the saturation properties of CO_2-based refrigerant blends without the need for extensive experimental data. Based on an ANN structure, the method used COSMO-RS molecular descriptors as input information and was trained based on dew and bubble points obtained with the polar soft-SAFT equation of state.

The initial phase of this study involved the pure-component characterization of 38 refrigerant compounds using polar soft-SAFT. This set included novel fluorine-based refrigerants and compounds with non-standard elements, such as R13I1 and RE170. The polar soft-SAFT EoS proved robust alignment with experimental data, maintaining vapor pressure and saturated densities within 3 % and 1 % deviations respectively, and estimating latent heat with a relative deviation below 5 % in all cases. Notably, compounds like R601a, R1224yd(E), R1234ze (Z) and R1336yf stood out due to exceptional precision in their characterization, with deviations below 1 % in refrigeration effect estimates. Next, the solubility profiles of CO_2-blends in selected refrigerants were characterized across multiple isotherms, using a temperature-independent binary parameter to closely align with experimental data. The model's robustness was further highlighted by its need for only minimal adjustments from ideality, finely tuning CO_2 solubility in twenty-seven refrigerant agents within a range of behaviors from purely azeotropic to highly zeotropic. Subsequently, a multitask ANN was trained using SAFT-derived characterizations as output targets, aiming to predict the saturation properties in both liquid and vapor phases of CO_2 refrigerant blends. Following validation by comparison with well-established thermodynamic models like NRTL, the analysis revealed an AARD below 3.5 % for both liquid and vapor phases in blends containing newly tested HFOs and HCFOs refrigerants. Within this scope, COSMO-RS σ-profiles were employed as molecular descriptors, effectively quantifying molecular interactions derived from quantum chemical calculations, providing data to be used in the absence of experimental measurements. The database included all pure-components and binary CO_2-based blends derived from preceding characterizations, presenting 531 distinct combinations for ANN training. The ANN's effectiveness was assessed with key performance indicators such as R^2, RMSE, AARD, SD_{av}, or NMAD along with residual analysis, revealing only 2.63 % and 2.44 % of total data as outliers for bubble and dew phase predictions, respectively. The ANN's fitting, consistent with SAFT targets across a broad range of mixtures and working conditions, showcases high reliability, wide-ranging applicability and transferability, evidenced by early outlier detection through residual analysis. Although high precision was maintained across datasets, the model exhibited a slightly lower performance in validation and testing subsets for dew phase estimations beyond 300 K, highlighting the importance of data diversity in ANN training stages.

In summary, this research combines artificial intelligence with thermodynamic modeling, enabling exploration of future refrigerant mixtures through computational approaches with proven predictive accuracy. The capability to predict and further validate saturation dew and bubble points of unmeasured systems enables the fitting of molecular-based coarse-grained models like the polar soft-SAFT, marking a significant improvement in thermodynamic predictions, reducing experimental dependency, and simplifying complex system modeling. This is a needed step to further evaluate the performance of these blends in cooling applications, for which there is limited to non-available experimental data. Given the versatility of the approach, future research should focus on extending the methodology to a broader range of refrigerant blends and properties of interest, while refining the ANN's structure and predictive capability. Collaboration with industry partners will be essential to integrate the approach into commercial refrigeration applications, ensuring its applicability in system design and optimization. Additionally, targeted experimental validation of CO_2-based mixtures with limited available data will further strengthen the model's reliability and expand its predictive scope.

CRediT authorship contribution statement

Albà Carlos G.: Writing – original draft, Methodology, Investigation. **Llovell Fèlix:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition. **Alkhatib Ismail I.I.:** Writing – review & editing, Validation, Investigation. **Vega Lourdes F.:** Writing – review & editing, 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

This work was done in the framework of projects NEW-F-Tech (Ref: TED2021–130959B-I00) and REFCICLA (Ref: PID2023–149713OB-I00) funded by MICIU/AEI/10.13039/501100011033/ and by the European Union NextGenerationEU/ PRTR. Additional support was provided by Khalifa University of Science and Technology through project RC2–2019-007 to the Research and Innovation Center on CO2 and Hydrogen (RICH Center) and through project RIG-2024–041. C. G. Albà acknowledges a FI-SDUR fellowship from the Catalan Government (AGAUR) and a visiting scholar grant for short stays provided by the RICH Center. Additional funding from AGAUR as a Consolidated Research Group (SGR 2021–00738) is appreciated. Finally, computational resources from the RICH Center and the Almesbar HPC at the Research Computing department at Khalifa University of Science and Technology are gratefully acknowledged.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jcou.2025.103072.

Data Availability

Supplementary information contains the ANN parameters to reproduce the resultsSupplem

References

[1] M.O. McLinden, M.L. Huber, R)Evolution of refrigerants, J. Chem. Eng. Data 65 (9) (2020) 4176–4193, https://doi.org/10.1021/acs.jced.0c00338.
[2] J.H. Koh, Z. Zakaria, Hydrocarbons as refrigerants–a review, ASEAN J. Sci. Technol. Dev. 34 (1) (2017) 35, https://doi.org/10.29037/ajstd.73.
[3] R. Ciconkov, Refrigerants: there is still no vision for sustainable solutions, Int. J. Refrig 86 (2018) 441–448, https://doi.org/10.1016/J.IJREFRIG.2017.12.006.
[4] G. Raabe, Molecular simulation studies on refrigerants past – present – future, Fluid Phase Equilib. 485 (2019) 190–198, https://doi.org/10.1016/j.fluid.2018.12.022.
[5] J.M. Calm, The next generation of refrigerants – historical review, considerations, and outlook, Int. J. Refrig 31 (7) (2008) 1123–1133, https://doi.org/10.1016/j.ijrefrig.2008.01.013.
[6] N. Abas, A.R. Kalair, N. Khan, A. Haider, Z. Saleem, M.S. Saleem, Natural and synthetic refrigerants, global warming: a review, Renew. Sustain. Energy Rev. 90 (2018) 557–569, https://doi.org/10.1016/j.rser.2018.03.099.
[7] J.M. Calm, D.A. Didion, Trade-offs in refrigerant selections: past, present, and future, Int. J. Refrig 21 (4) (1998) 308–321, https://doi.org/10.1016/S0140-7007(97)00089-3.
[8] Z. Yang, B. Feng, H. Ma, L. Zhang, C. Duan, B. Liu, Y. Zhang, S. Chen, Z. Yang, Analysis of lower GWP and flammable alternative refrigerants, Int. J. Refrig 126 (2021) 12–22, https://doi.org/10.1016/j.ijrefrig.2021.01.022.
[9] V. Nair, HFO refrigerants: a review of present status and future prospects, Int. J. Refrig 122 (2021) 156–170, https://doi.org/10.1016/j.ijrefrig.2020.10.039.
[10] UN Environment Ozone Secretariat, Ratification of the Kigali Amendment. United Nations Environ. Program, 2017 (No. February).
[11] B.K. Sovacool, S. Griffiths, J. Kim, M. Bazilian, Climate change and industrial F-gases: a critical and systematic review of developments, sociotechnical systems

and policy options for reducing synthetic greenhouse gas Emissions, Renew. Sustain. Energy Rev. 141 (2021) 110759, https://doi.org/10.1016/j.rser.2021.110759.
[12] AIM Act | US EPA
[13] M. Ghodbane, An investigation of R152a and hydrocarbon refrigerants in mobile air conditioning, L150/195-230, Off. J. Eur. Union 2014 (1999), https://doi.org/10.4271/1999-01-0874.
[14] Mclinden, M.O.; Kazakov, A.F.; Brown, P.J.S.; Brignoli, R.; Bell, I.H.; Domanski, P. A. Options for Low-GWP Refrigerants in Small Air-Conditioning Systems. 2019, No. November, 1–18.
[15] C.G. Albà, I.I.I. Alkhatib, F. Llovell, L.F. Vega, Hunting sustainable refrigerants fulfilling technical, environmental, safety and economic requirements, Renew. Sustain. Energy Rev. 188 (September) (2023) 113806, https://doi.org/10.1016/j.rser.2023.113806.
[16] A. Mota-Babiloni, J. Navarro-Esbrí, P. Makhnatch, F. Molés, Refrigerant R32 as lower GWP working fluid in residential air conditioning systems in Europe and the USA, Renew. Sustain. Energy Rev. (December 2017) 1031–1042, https://doi.org/10.1016/j.rser.2017.05.216.
[17] S. Kujak, K. Schultz, Insights into the next Generation HVAC&R Refrigerant Future, Sci. Technol. Built Environ. 22 (8) (2016) 1226–1237, https://doi.org/10.1080/23744731.2016.1203239.
[18] S. Yadav, J. Liu, S.C. Kim, A comprehensive study on 21st-century refrigerants - R290 and R1234yf: a review, Int. J. Heat. Mass Transf. 182 (2022) 121947, https://doi.org/10.1016/j.ijheatmasstransfer.2021.121947.
[19] A. Miyara, Developments of next Generation refrigerants and heat transfer, MATEC Web Conf. 204 (2018) 00004, https://doi.org/10.1051/matecconf/201820400004.
[20] D. Sánchez, A. Andreu-Nácher, D. Calleja-Anta, R. Llopis, R. Cabello, Energy impact evaluation of different low-GWP alternatives to replace R134a in a beverage cooler. experimental analysis and optimization for the pure refrigerants R152a, R1234yf, R290, R1270, R600a and R744, Energy Convers. Manag 256 (2022) 115388, https://doi.org/10.1016/j.enconman.2022.115388.
[21] D. Sánchez, R. Cabello, R. Llopis, I. Arauzo, J. Catalán-Gil, E. Torrella, Energy performance evaluation of R1234yf, R1234ze(E), R600a, R290 and R152a as Low-GWP R134a alternatives, Int. J. Refrig 74 (2017) 269–282, https://doi.org/10.1016/j.ijrefrig.2016.09.020.
[22] N. Kumma, S.H. Kruthiventi, Current status of refrigerants used in domestic applications: a review, Renew. Sustain. Energy Rev. 189 (2024) 114073, https://doi.org/10.1016/j.rser.2023.114073.
[23] L. Fedele, G. Lombardo, I. Greselin, D. Menegazzo, S. Bobbo, Thermophysical properties of low gwp refrigerants: an update, Int. J. Thermophys. 44 (5) (2023) 80, https://doi.org/10.1007/s10765-023-03191-5.
[24] P. Bansal, A Review – Status of CO2 as a low temperature refrigerant: fundamentals and R&D opportunities, Appl. Therm. Eng. 41 (2012) 18–29, https://doi.org/10.1016/j.applthermaleng.2011.12.006.
[25] F. Bruno, M. Belusko, E. Halawa, CO2 Refrigeration and heat pump systems—a comprehensive review, Energies 12 (15) (2019) 2959, https://doi.org/10.3390/en12152959.
[26] K. Zolcer Skačanová, M. Battesti, Global market and policy trends for CO2 in refrigeration, Int. J. Refrig 107 (2019) 98–104, https://doi.org/10.1016/j.ijrefrig.2019.08.010.
[27] Y. Heredia-Aricapa, J.M. Belman-Flores, A. Mota-Babiloni, J. Serrano-Arellano, J. J. García-Pabón, Overview of low GWP mixtures for the replacement of HFC refrigerants: R134a, R404A and R410A, Int. J. Refrig 111 (2020) 113–123, https://doi.org/10.1016/j.ijrefrig.2019.11.012.
[28] K. Braimakis, A. Mikelis, A. Charalampidis, S. Karellas, Exergetic performance of CO2 and ultra-low gwp refrigerant mixtures as working fluids in ORC for waste heat recovery, Energy 203 (2020) 117801, https://doi.org/10.1016/j.energy.2020.117801.
[29] B. Dai, M. Li, Y. Ma, Thermodynamic analysis of carbon dioxide blends with Low GWP (Global Warming Potential) working fluids-based transcritical rankine cycles for low-grade heat energy recovery, Energy 64 (2014) 942–952, https://doi.org/10.1016/j.energy.2013.11.019.
[30] B.O. Bolaji, Theoretical assessment of new low global warming potential refrigerant mixtures as eco-friendly alternatives in domestic refrigeration Systems, Sci. Afr. 10 (2020) e00632, https://doi.org/10.1016/j.sciaf.2020.e00632.
[31] B. Yu, H. Ouyang, J. SHI, W. LIU, J. CHEN, Evaluation of Low-GWP and mildly flammable mixtures as new alternatives for R410A in air-conditioning and heat pump system, Int. J. Refrig 121 (2021) 95–104, https://doi.org/10.1016/j.ijrefrig.2020.09.018.
[32] G. Vaccaro, A. Milazzo, L. Talluri, Thermodynamic assessment of trans-critical refrigeration systems utilizing CO2-based mixtures, Int. J. Refrig 147 (2023) 61–70, https://doi.org/10.1016/j.ijrefrig.2022.09.013.
[33] D. Sánchez, F. Vidan-Falomir, L. Nebot-Andrés, R. Llopis, R. Cabello, Alternative blends of CO2 for transcritical refrigeration systems. experimental approach and energy analysis, Energy Convers. Manag 279 (2023), https://doi.org/10.1016/J.ENCONMAN.2023.116690.
[34] E.W. Lemmon, M.L. Huber, M.O. McLinden, NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport Properties (REFPROP), Version 9.1, National Institute of Standards and Technology, Standard Reference Data Program, 2010.
[35] G. Raabe, Molecular simulation data for the vapor-liquid phase equilibria of binary mixtures of HFO-1123 with R-32, R-1234yf, R-1234ze(E), R-134a and CO2 and Their Modelling by the PCP-SAFT equation of state, Data Br. 25 (2019) 104014, https://doi.org/10.1016/j.dib.2019.104014.
[36] M.S. Sadaghiani, A. Arami-Niya, B. Marsh, S.Z.S. Al Ghafri, E.F. May, Vapor–liquid equilibria for carbon dioxide + 3,3,3-trifluoropropene binary mixtures at temperatures between (288 and 348) K, J. Chem. Eng. Data 66 (11) (2021) 4044–4055, https://doi.org/10.1021/acs.jced.1c00297.
[37] H. Matsuda, T. Suga, T. Tsuji, K. Tochigi, K. Kurihara, A.K. Nelson, C. McCabe, Vapor-liquid equilibria for binary systems carbon dioxide + 1,1,1,2,3,3-Hexafluoro-3-(2,2,2-Trifluoroethoxy)Propane or 1-Ethoxy-1,1,2,2,3,3,4,4,4-nonafluorobutane at 303.15–323.15 K, Fluid Phase Equilib. 524 (2020) 112814, https://doi.org/10.1016/j.fluid.2020.112814.
[38] J. Luo, Z. Ye, Z. Zhao, K. Yang, S. Zhang, Q. Wang, Vapor-liquid equilibrium measurement and heating performance modeling on eco-friendly zeotropic blends of CO2/R1234ze(Z) and CO2/R1336mzz(E), Appl. Therm. Eng. 229 (2023) 120576, https://doi.org/10.1016/j.applthermaleng.2023.120576.
[39] J. El Abbadi, J. Brocus, A. Valtz, C. Coquelet, C. Houriez, Experimental measurements and modeling of vapor–liquid equilibria for eight mixtures containing trans -1-Chloro-3,3,3-Trifluoropropene (R1233zd(E)) and 2-Chloro-3,3,3-Trifluoropropene (R1233xf), J. Chem. Eng. Data 68 (9) (2023) 2316–2331, https://doi.org/10.1021/acs.jced.3c00296.
[40] I.H. Bell, D. Riccardi, A. Bazyleva, M.O. McLinden, Survey of data and models for refrigerant mixtures containing halogenated olefins, J. Chem. Eng. Data 66 (6) (2021) 2335–2354, https://doi.org/10.1021/acs.jced.1c00192.
[41] X. Yang, A. Arami-Niya, X. Xiao, D. Kim, S.Z.S. Al Ghafri, T. Tsuji, Y. Tanaka, Y. Seiki, E.F. May, Viscosity measurements of binary and multicomponent refrigerant mixtures containing HFC-32, HFC-125, HFC-134a, HFO-1234yf, and CO 2, J. Chem. Eng. Data 65 (9) (2020) 4252–4262, https://doi.org/10.1021/acs.jced.0c00228.
[42] D. Kim, X. Yang, A. Arami-Niya, D. Rowland, X. Xiao, S.Z.S. Al Ghafri, T. Tsuji, Y. Tanaka, Y. Seiki, E.F. May, Thermal conductivity measurements of refrigerant mixtures containing hydrofluorocarbons (HFC-32, HFC-125, HFC-134a), hydrofluoroolefins (HFO-1234yf), and carbon dioxide (CO2), J. Chem. Thermodyn. 151 (2020) 106248, https://doi.org/10.1016/j.jct.2020.106248.
[43] X. Yang, D. Kim, E.F. May, I.H. Bell, Entropy scaling of thermal conductivity: application to refrigerants and their mixtures, Ind. Eng. Chem. Res. 60 (35) (2021) 13052–13070, https://doi.org/10.1021/ACS.IECR.1C02154/SUPPL_FILE/IE1C02154_SI_001.PDF.
[44] I.I.I. Alkhatib, F. Llovell, L.F. Vega, Assessing the effect of impurities on the thermophysical properties of methane-based energy systems using polar soft-SAFT, Fluid Phase Equilib. 527 (2021) 112841, https://doi.org/10.1016/j.fluid.2020.112841.
[45] I.I.I. Alkhatib, A. AlHajaj, A. Almansoori, L.F. Vega, Accurate predictions of the effect of hydrogen composition on the thermodynamics and transport properties of natural gas, Ind. Eng. Chem. Res. 61 (18) (2022) 6214–6234, https://doi.org/10.1021/acs.iecr.2c00363.
[46] I.I.I. Alkhatib, O. Khalifa, D. Bahamon, M.R.M. Abu-Zahra, L.F. Vega, Sustainability criteria as a game changer in the search for hybrid solvents for CO2 and H2S Removal, Sep. Purif. Technol. 277 (2021) 119516, https://doi.org/10.1016/j.seppur.2021.119516.
[47] I.I.I. Alkhatib, A. Galindo, L.F. Vega, Systematic STudy of the Effect of the Co-solvent on the performance of amine-based solvents for CO2 Capture, Sep. Purif. Technol. 282 (2022) 120093, https://doi.org/10.1016/j.seppur.2021.120093.
[48] I.I.I. Alkhatib, D. Bahamon, A. Al Hajaj, L.F. Vega, Molecular thermodynamic modeling of hybrid ionic liquids for biogas upgrading, Ind. Eng. Chem. Res. 61 (33) (2022) 12190–12207, https://doi.org/10.1021/acs.iecr.2c00710.
[49] C.G. Albà, I.I.I. Alkhatib, F. Llovell, L.F. Vega, Assessment of low global warming potential refrigerants for drop-in replacement by connecting their molecular features to their performance, ACS Sustain. Chem. Eng. 9 (50) (2021) 17034–17048, https://doi.org/10.1021/acssuschemeng.1c05985.
[50] N.L. Mai, Y.-M. Koo, Quantitative prediction of lipase reaction in ionic liquids by QSAR Using COSMO-RS molecular descriptors, Biochem. Eng. J. 87 (2014) 33–40, https://doi.org/10.1016/j.bej.2014.03.010.
[51] I. Díaz, M. Rodríguez, M. González-Miquel, E.J. González, COSMO-derived descriptors applied in ionic liquids physical property modelling using machine learning algorithms, Comput. Aided Chem. Eng. 43 (2018) 121–126, https://doi.org/10.1016/B978-0-444-64235-6.50023-1.
[52] T. Lemaoui, A.S. Darwish, G. Almustafa, A. Boublia, P.R. Sarika, N.A. Jabbar, T. Ibrahim, P. Nancarrow, K.K. Yadav, A.M. Fallatah, et al., Machine learning approach to map the thermal conductivity of over 2,000 neoteric solvents for green energy storage applications, Energy Storage Mater. 59 (April) (2023) 102795, https://doi.org/10.1016/j.ensm.2023.102795.
[53] M. Panić, M. Radović, M. Cvjetko Bubalo, K. Radošević, M. Rogošić, J.A. P. Coutinho, I. Radojčić Redovniković, A. Jurinjak Tušek, Prediction of PH value of aqueous acidic and basic deep eutectic solvent using COSMO-RS σ profiles' molecular descriptors, Molecules 27 (14) (2022) 4489, https://doi.org/10.3390/molecules27144489.
[54] T. Lemaoui, A. Boublia, A.S. Darwish, M. Alam, S. Park, B. Jeon, F. Banat, Y. Benguerba, I.M. AlNashef, Predicting the surface tension of deep eutectic solvents using artificial neural networks, ACS Omega 7 (36) (2022) 32194–32207, https://doi.org/10.1021/acsomega.2c03458.
[55] A. Boublia, T. Lemaoui, F. Abu Hatab, A.S. Darwish, F. Banat, Y. Benguerba, I. M. AlNashef, Molecular-based artificial neural network for predicting the electrical conductivity of deep eutectic solvents, J. Mol. Liq. 366 (2022) 120225, https://doi.org/10.1016/j.molliq.2022.120225.
[56] T. Lemaoui, A. Boublia, S. Lemaoui, A.S. Darwish, B. Ernst, M. Alam, Y. Benguerba, F. Banat, I.M. AlNashef, Predicting the CO 2 capture capability of deep eutectic solvents and screening over 1000 of their combinations using

machine learning, ACS Sustain. Chem. Eng. 11 (26) (2023) 9564–9580, https://doi.org/10.1021/acssuschemeng.3c00415.
[57] A.S. Darwish, R. Abu Alwan, A. Boublia, T. Lemaoui, Y. BENGUERBA, I. Al Nashef, F. Banat, Accurate heat capacity predictions of deep eutectic solvents using machine learning, J. Meas. (2023), https://doi.org/10.2139/ssrn.4458944.
[58] T. Lemaoui, F. Abu Hatab, A.S. Darwish, A. Attoui, N.E.H. Hammoudi, G. Almustafa, M. Benaicha, Y. Benguerba, I.M. Alnashef, Molecular-based guide to predict the ph of eutectic solvents: promoting an efficient design approach for new green solvents, ACS Sustain. Chem. Eng. 9 (17) (2021) 5783–5808, https://doi.org/10.1021/acssuschemeng.0c07367.
[59] A. Sosa, J. Ortega, L. Fernández, J. Palomar, Development of a method to model the mixing energy of solutions using COSMO molecular descriptors linked with a semi-empirical model using a combined ANN-QSPR Methodology, Chem. Eng. Sci. 224 (2020) 115764, https://doi.org/10.1016/j.ces.2020.115764.
[60] C.G. Albà, I.I.I. Alkhatib, L.F. Vega, F. Llovell, Mapping the FLammability Space of Sustainable Refrigerant Mixtures through an Artificial Neural Network Based on Molecular Descriptors, ACS Sustain. Chem. Eng. 12 (31) (2024) 11561–11577, https://doi.org/10.1021/acssuschemeng.4c01961.
[61] A. Boublia, T. Lemaoui, J. AlYammahi, A.S. Darwish, A. Ahmad, M. Alam, F. Banat, Y. Benguerba, I.M. AlNashef, Multitask neural network for mapping the glass transition and melting temperature space of homo- and co-polyhydroxyalkanoates using σ profiles molecular inputs, ACS Sustain. Chem. Eng. 11 (1) (2023) 208–227, https://doi.org/10.1021/acssuschemeng.2c05225.
[62] J. Wang, Z. Song, L. Chen, T. Xu, L. Deng, Z. Qi, Prediction of CO2 solubility in deep eutectic solvents using random forest model based on COSMO-RS-derived descriptors, Green. Chem. Eng. 2 (4) (2021) 431–440, https://doi.org/10.1016/j.gce.2021.08.002.
[63] C.Y. Chen, L.H. Wang, C.M. Hsieh, S.T. Lin, Prediction of solid-liquid-gas equilibrium for binary mixtures of carbon dioxide + organic compounds from approaches based on the COSMO-SAC Model, J. Supercrit. Fluids 133 (2018) 318–329, https://doi.org/10.1016/J.SUPFLU.2017.08.008.
[64] S. Deng, W. Su, L. Zhao, A neural network for predicting normal boiling point of pure refrigerants using molecular groups and a topological index, Int. J. Refrig 63 (2016) 63–71, https://doi.org/10.1016/j.ijrefrig.2015.10.025.
[65] B. Liu, M. Karimi Nouroddin, Application of artificial intelligent approach to predict the normal boiling point of refrigerants, Int. J. Chem. Eng. 2023 (2023) 1–9, https://doi.org/10.1155/2023/6809569.
[66] G. Wang, P. Hu, Q. Chen, C. Xu, J. Zhao, Prediction of critical temperature of binary refrigerant mixtures by neural network, Int. J. Refrig 161 (2024) 210–220, https://doi.org/10.1016/j.ijrefrig.2024.02.031.
[67] I.I.I. Alkhatib, C.G. Albà, A.S. Darwish, F. Llovell, L.F. Vega, Searching for sustainable refrigerants by bridging molecular modeling with machine learning, Ind. Eng. Chem. Res. 61 (21) (2022) 7414–7429, https://doi.org/10.1021/acs.iecr.2c00719.
[68] I.I.I. Alkhatib, L.M.C. Pereira, J. Torne, L.F. Vega, Polar Soft-SAFT: theory and comparison with molecular simulations and experimental data of pure polar fluids, Phys. Chem. Chem. Phys. 22 (23) (2020) 13171–13191, https://doi.org/10.1039/D0CP00846J.
[69] J.C. Pàmies, L.F. Vega, Vapor−liquid equilibria and critical behavior of heavy n -alkanes using transferable parameters from the Soft-SAFT equation of state, Ind. Eng. Chem. Res. 40 (11) (2001) 2532–2543, https://doi.org/10.1021/ie000944x.
[70] M.S. Wertheim, Fluids with highly directional attractive forces. I. Statistical thermodynamics, J. Stat. Phys. 35 (1–2) (1984) 19–34, https://doi.org/10.1007/BF01017362.
[71] M.S. Wertheim, Fluids with highly directional attractive forces. II. thermodynamic perturbation theory and integral equations, J. Stat. Phys. 35 (1–2) (1984) 35–47, https://doi.org/10.1007/BF01017363.
[72] M.S. Wertheim, Fluids with highly directional attractive forces. III. Multiple attraction sites, J. Stat. Phys. 42 (3–4) (1986) 459–476, https://doi.org/10.1007/BF01127721.
[73] M.S. Wertheim, Fluids with highly directional attractive forces. IV. Equilibrium polymerization, J. Stat. Phys. 42 (3–4) (1986) 477–492, https://doi.org/10.1007/BF01127722.
[74] F.J. Blas, L.F. Vega, Thermodynamic behaviour of homonuclear and heteronuclear lennard-jones chains with association sites from simulation and theory, Mol. Phys. 92 (1) (1997) 135–150, https://doi.org/10.1080/002689797170707.
[75] J.K. Johnson, J.A. Zollweg, K.E. Gubbins, The lennard-jones equation of state revisited, Mol. Phys. 78 (3) (1993) 591–618, https://doi.org/10.1080/00268979300100411.
[76] W.G. Chapman, K.E. Gubbins, G. Jackson, M. Radosz, SAFT: equation-of-state solution model for associating fluids, Fluid Phase Equilib. 52 (1989) 31–38, https://doi.org/10.1016/0378-3812(89)80308-5.
[77] W.G. Chapman, K.E. Gubbins, G. Jackson, M. Radosz, New reference equation of state for associating liquids, Ind. Eng. Chem. Res. 29 (8) (1990) 1709–1721, https://doi.org/10.1021/ie00104a021.
[78] M. Luckas, K. Lucas, U. Deiters, K.E. Gubbins, Integrals over pair- and triplet-correlation functions for the Lennard-Jones (12-6)-Fluid, Mol. Phys. 57 (2) (1986) 241–253, https://doi.org/10.1080/00268978600100191.
[79] K.E. Gubbins, C.H. Twu, Thermodynamics of polyatomic fluid mixtures-I: theory, Chem. Eng. Sci. 33 (7) (1978) 863–878, https://doi.org/10.1016/0009-2509(78)85176-8.
[80] C.H. Twu, K.E. Gubbins, Thermodynamics of polyatomic fluid mixtures—II: Polar, quadrupolar and octopolar molecules, Chem. Eng. Sci. 33 (7) (1978) 879–887, https://doi.org/10.1016/0009-2509(78)85177-x.
[81] P.K. Jog, W.G. Chapman, Application of Wertheim's thermodynamic perturbation theory to dipolar hard sphere chains, Mol. Phys. 97 (3) (1999) 307–319, https://doi.org/10.1080/00268970601076467.
[82] P.K. Jog, S.G. Sauer, J. Blaesing, W.G. Chapman, Application of dipolar chain theory to the phase behavior of polar fluids and mixtures, Ind. Eng. Chem. Res. 40 (21) (2001) 4641–4648, https://doi.org/10.1021/ie010264+.
[83] G. Stell, J.C. Rasaiah, H. Narang, Thermodynamic perturbation theory for simple polar fluids. II, Mol. Phys. 27 (5) (1974) 1393–1414, https://doi.org/10.1080/00268977400101181.
[84] I.I.I. Alkhatib, L.F. Vega, Quantifying the effect of polarity on the behavior of mixtures of n-alkanes with dipolar solvents using polar soft-statistical associating fluid theory (Polar Soft-SAFT), AIChE J. 67 (3) (2021), https://doi.org/10.1002/aic.16649.
[85] C.G. Albà, F. Llovell, L.F. Vega, Searching for suitable lubricants for low global warming potential refrigerant R513A using molecular-based models: solubility and performance in refrigeration cycles, Int. J. Refrig 128 (2021) 252–263, https://doi.org/10.1016/j.ijrefrig.2021.04.010.
[86] I.I.I. Alkhatib, L.F. Vega, Quantifying the Effect of polar interactions on the behavior of binary mixtures: phase, interfacial, and excess properties, J. Chem. Phys. 154 (16) (2021) 164503, https://doi.org/10.1063/5.0046034.
[87] C.G. Albà, I.I.I. Alkhatib, F. Llovell, L.F. Vega, A novel approach for designing efficient and sustainable cooling cycles with low global warming potential refrigerants, Appl. Therm. Eng. 246 (2024) 122895, https://doi.org/10.1016/j.applthermaleng.2024.122895.
[88] D. Macgowan, J.L. Lebowitz, E.M. Waisman, Van Der Waals one-fluid theory: justification and generalisation, Chem. Phys. Lett. 114 (3) (1985) 321–324, https://doi.org/10.1016/0009-2614(85)80923-4.
[89] H. Madani, A. Valtz, C. Coquelet, A.H. Meniai, D. Richon, Vapor+liquid) Equilibrium Data for (Carbon Dioxide+1,1-Difluoroethane) System at Temperatures from (258 to 343)K and Pressures up to about 8MPa, J. Chem. Thermodyn. 40 (10) (2008) 1490–1494, https://doi.org/10.1016/j.jct.2008.06.002.
[90] C. Duran-Valencia, G. Pointurier, A. Valtz, P. Guilbot, D. Richon, Vapor−Liquid Equilibrium (VLE) data for the carbon dioxide (CO 2) + 1,1,1,2-tetrafluoroethane (R134a) system at temperatures from 252.95 K to 292.95 K and pressures up to 2 MPa, J. Chem. Eng. Data 47 (1) (2002) 59–61, https://doi.org/10.1021/je010075y.
[91] N. Juntarachat, A. Valtz, C. Coquelet, R. Privat, J.-N. Jaubert, Experimental measurements and correlation of vapor–liquid equilibrium and critical data for the CO 2 + R1234yf and CO 2 + R1234ze(E) binary mixtures, Int. J. Refrig 47 (2014) 141–152, https://doi.org/10.1016/j.ijrefrig.2014.09.001.
[92] S. Wang, R. Fauve, C. Coquelet, A. Valtz, C. Houriez, P.-A. Artola, E.El Ahmar, B. Rousseau, H. Hu, Vapor–Liquid equilibrium and molecular simulation data for carbon dioxide (CO2) + Trans-1,3,3,3-Tetrafluoroprop-1-Ene (R-1234ze(E)) mixture at temperatures from 283.32 to 353.02 K and pressures up to 7.6 MPa, Int. J. Refrig 98 (2019) 362–371, https://doi.org/10.1016/j.ijrefrig.2018.10.032.
[93] Z. Yuan, Y. Tu, C. Wang, Y. Zhao, X. Dong, Experimental Research on (Vapor + Liquid) Equilibria for the {trifluoroiodomethane (CF3I) + Carbon Dioxide (CO2)} System from 243.150 to 273.150 K, J. Chem. Thermodyn. 101 (2016) 49–53, https://doi.org/10.1016/j.jct.2016.05.012.
[94] A. Klamt, G. Schüürmann, COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient, J. Chem. Soc. Perkin Trans. 2 (5) (1993) 799–805, https://doi.org/10.1039/P29930000799.
[95] E. Mullins, R. Oldland, Y.A. Liu, S. Wang, S.I. Sandler, C.-C. Chen, M. Zwolak, K. C. Seavey, Sigma-profile database for using COSMO-based thermodynamic methods, Ind. Eng. Chem. Res. 45 (12) (2006) 4389–4415, https://doi.org/10.1021/ie060370h.
[96] L. Moity, M. Durand, A. Benazzouz, C. Pierlot, V. Molinier, J.-M. Aubry, Panorama of sustainable solvents using the COSMO-RS approach, Green. Chem. 14 (4) (2012) 1132, https://doi.org/10.1039/c2gc16515e.
[97] T. Lemaoui, A.S. Darwish, A. Attoui, F. Abu Hatab, N.E.H. Hammoudi, Y. Benguerba, L.F. Vega, I.M. Alnashef, Predicting the density and viscosity of hydrophobic eutectic solvents: towards the development of sustainable solvents, Green. Chem. 22 (23) (2020) 8511–8530, https://doi.org/10.1039/D0GC03077E.
[98] J.S. Torrecilla, J. Palomar, J. Lemus, F. Rodríguez, A quantum-chemical-based guide to analyze/quantify the cytotoxicity of ionic liquids, Green. Chem. 12 (1) (2010) 123–134, https://doi.org/10.1039/B919806G.
[99] Y. Benguerba, I.M. Alnashef, A. Erto, M. Balsamo, B. Ernst, A quantitative prediction of the viscosity of amine based DESs Using Sσ-profile molecular descriptors, J. Mol. Struct. 1184 (2019) 357–363, https://doi.org/10.1016/j.molstruc.2019.02.052.
[100] Y. Zhao, Y. Huang, X. Zhang, S. Zhang, A quantitative prediction of the viscosity of ionic liquids using S σ-profile molecular descriptors, Phys. Chem. Chem. Phys. 17 (5) (2015) 3761–3767, https://doi.org/10.1039/C4CP04712E.
[101] T. Lemaoui, A.S. Darwish, N.E.H. Hammoudi, F. Abu Hatab, A. Attoui, I. M. Alnashef, Y. Benguerba, Prediction of electrical conductivity of deep eutectic solvents using COSMO-RS sigma profiles as molecular descriptors: a quantitative structure–property relationship study, Ind. Eng. Chem. Res. 59 (29) (2020) 13343–13354, https://doi.org/10.1021/acs.iecr.0c02542.
[102] S. Haykin, Neural Networks and Learning Machines, 3/E, Pearson Education India, 2009.
[103] F. Marini, R. Bucci, A.L. Magrì, A.D. Magrì, Artificial neural networks in chemometrics: history, examples and perspectives, Microchem. J. 88 (2) (2008) 178–185, https://doi.org/10.1016/j.microc.2007.11.008.

[104] M. Jalali-Heravi, A. Kyani, Use of computer-assisted methods for the modeling of the retention time of a variety of volatile organic compounds: a PCA-MLR-ANN approach, J. Chem. Inf. Comput. Sci. 44 (4) (2004) 1328–1335, https://doi.org/10.1021/ci0342270.
[105] R. Kocjančič, J. Zupan, Application of a feed-forward artificial neural network as a mapping device, J. Chem. Inf. Comput. Sci. 37 (6) (1997) 985–989, https://doi.org/10.1021/ci970223h.
[106] K. Paduszyński, U. Domańska, Viscosity of ionic liquids: an extensive database and a new group contribution model based on a feed-forward artificial neural network, J. Chem. Inf. Model 54 (5) (2014) 1311–1324, https://doi.org/10.1021/ci500206u.
[107] M. Karelson, D.A. Dobchev, O.V. Kulshyn, A.R. Katritzky, Neural networks convergence using physicochemical data, J. Chem. Inf. Model 46 (5) (2006) 1891–1897, https://doi.org/10.1021/ci0600206.
[108] A.L. Nazarova, L. Yang, K. Liu, A. Mishra, R.K. Kalia, K. Nomura, A. Nakano, P. Vashishta, P. Rajak, Dielectric polymer property prediction using recurrent neural networks with optimizations, J. Chem. Inf. Model 61 (5) (2021) 2175–2186, https://doi.org/10.1021/acs.jcim.0c01366.
[109] M. Abdi-Khanghah, A. Bemani, Z. Naserzadeh, Z. Zhang, Prediction of solubility of N-alkanes in supercritical CO2 using RBF-ANN and MLP-ANN, J. CO2 Util. 25 (2018) 108–119, https://doi.org/10.1016/j.jcou.2018.03.008.
[110] Jain, P.; Meenu. Recognition of Mechanical Tools Using Artificial Neural Network; 2021; pp 637–644. https://doi.org/10.1007/978-981-15-8704-7_78.
[111] K. Wu, G.-W. Wei, Quantitative toxicity prediction using topology based multitask deep neural networks, J. Chem. Inf. Model 58 (2) (2018) 520–531, https://doi.org/10.1021/acs.jcim.7b00558.
[112] C. Ji, S. Yuan, Z. Jiao, M. Huffman, M.M. El-Halwagi, Q. Wang, Predicting flammability-leading properties for liquid aerosol safety via machine learning, Process Saf. Environ. Prot. 148 (2021) 1357–1366, https://doi.org/10.1016/j.psep.2021.03.012.
[113] G. Di Nicola, G. Coccia, M. Pierantozzi, S. Tomassetti, Equations for the surface tension of low GWP halogenated alkene refrigerants and their blends, Int. J. Refrig 86 (2018) 410–421, https://doi.org/10.1016/j.ijrefrig.2017.11.023.
[114] M.T. Bowers, J.B. Laudenslager, Mechanism of charge transfer reactions: reactions of rare gas ions with the trans-, Cis-, and 1,1-Difluoroethylene geometric isomers, J. Chem. Phys. 56 (9) (1972) 4711–4712, https://doi.org/10.1063/1.1677922.
[115] G. Raabe, Purely predictive vapor–liquid equilibrium properties of 3,3,4,4,4-pentafluoro-1-Butene (HFO-1345fz), 2,3,3,4,4,4-Hexafluoro-1-Butene (HFO-1336yf), and Trans -1-Chloro-2,3,3,3-Tetrafluoropropene (HCFO-1224yd(E)) from molecular simulation, J. Chem. Eng. Data 65 (9) (2020) 4318–4325, https://doi.org/10.1021/acs.jced.0c00325.
[116] F. Llovell, C.J. Peters, L.F. Vega, Second-order thermodynamic derivative properties of selected mixtures by the Soft-SAFT equation of state, Fluid Phase Equilib. 248 (2) (2006) 115–122, https://doi.org/10.1016/j.fluid.2006.07.018.
[117] F.J. Blas, L.F. Vega, Prediction of binary and ternary diagrams using the statistical associating fluid theory (SAFT) equation of state, Ind. Eng. Chem. Res. 5885 (1996) (1998) 660–674, https://doi.org/10.1021/ie970449+.
[118] D. Jovell, R. Gonzalez-Olmos, F. Llovell, A computational drop-in assessment of hydrofluoroethers in organic rankine cycles, Energy 254 (2022) 124319, https://doi.org/10.1016/j.energy.2022.124319.
[119] R. Low, Evaluation of potential use of R1132A as a refrigerant blend component (p Paper), Proc. 1st IIR Int. Conf. Appl. HFO Refrig. (2018) 1183, https://doi.org/10.18462/iir.hfo.2018.1183.
[120] U.A. Perera, N. Sakoda, T. Miyazaki, K. Thu, Y. Higashi, Measurements of saturation pressures for the novel refrigerant R1132(E), Int. J. Refrig 135 (2022) 148–153, https://doi.org/10.1016/j.ijrefrig.2021.12.014.
[121] C. Kondou, Okumura, Miura, Liu, Molecular simulation for liquid vapour equilibria of ethylene,R1132a, R1132(E) and R1123, 30028776 (2021) 1006, https://doi.org/10.18462/iir.HFO.2021.1006.
[122] N. Sakoda, Y. Higashi, R. Akasaka, Measurements of PvT properties, vapor pressures, saturated densities, and critical parameters for trans -1,1,1,4,4,4-Hexafluoro-2-Butene (R1336mzz(E)), J. Chem. Eng. Data 66 (1) (2021) 734–739, https://doi.org/10.1021/acs.jced.0c00848.
[123] X. Zhang, G. Zhao, J. Yin, S. Ma, Experimental investigation of saturated liquid kinematic viscosity and surface tension of two isomeric refrigerants Trans-1,1,1,4,4,4-Hexafluoro-Butene (R1336mzz(E)) and Cis-1,1,1,4,4,4-Hexafluoro-Butene (R1336mzz(Z)) by surface light scattering, Fluid Phase Equilib. 559 (2022) 113468, https://doi.org/10.1016/j.fluid.2022.113468.
[124] Y.-Y. Duan, M.-S. Zhu, L.-Z. Han, Experimental vapor pressure data and a vapor pressure equation for trifluoroiodomethane (CF3I), Fluid Phase Equilib. 121 (1–2) (1996) 227–234, https://doi.org/10.1016/0378-3812(96)03005-1.
[125] Y.-Y. Duan, L. Shi, M.-S. Zhu, L.-Z. Han, Critical parameters and saturated density of trifluoroiodomethane (CF 3 I), J. Chem. Eng. Data 44 (3) (1999) 501–504, https://doi.org/10.1021/je980251b.
[126] F. Llovell, J.C. Pàmies, L.F. Vega, Thermodynamic properties of lennard-jones chain molecules: renormalization-group corrections to a modified statistical associating fluid theory, J. Chem. Phys. 121 (21) (2004) 10715–10724, https://doi.org/10.1063/1.1809112.
[127] M. Gong, K. Cheng, X. Dong, H. Guo, Y. Zhao, J. Wu, Measurements of Isothermal (Vapor+liquid) Phase Equilibrium for {trifluoroiodomethane (R13I1)+ 1,1-Difluoroethane (R152a)} from T= (258.150 to 283.150)K, J. Chem. Thermodyn. 88 (2015) 90–95, https://doi.org/10.1016/j.jct.2015.04.011.
[128] N.A. Menad, A. Hemmati-Sarapardeh, A. Varamesh, S. Shamshirband, Predicting solubility of CO2 in brine by advanced machine learning systems: application to carbon capture and sequestration, J. CO2 Util. 33 (2019) 83–95, https://doi.org/10.1016/j.jcou.2019.05.009.
[129] M. Mesbah, S. Shahsavari, E. Soroush, N. Rahaei, M. Rezakazemi, Accurate prediction of miscibility of CO2 and supercritical CO2 in ionic liquids using machine learning, J. CO2 Util. 25 (2018) 99–107, https://doi.org/10.1016/j.jcou.2018.03.004.