



OPEN

On the evaluation of the carbon dioxide solubility in polymers using gene expression programming

Behnam Amiri-Ramsheh¹, Menad Nait Amar², Mohammadhadi Shateri³✉ & Abdolhossein Hemmati-Sarapardeh^{1,4}✉

Evaluation, prediction, and measurement of carbon dioxide (CO₂) solubility in different polymers are crucial for engineers in various chemical applications, such as extraction and generation of novel materials. In this paper, correlations based on gene expression programming (GEP) were generated to predict the value of carbon dioxide solubility in three polymers. Results showed that the generated correlations could represent an outstanding efficiency and provide predictions for carbon dioxide solubility with satisfactory average absolute relative errors of 9.71%, 5.87%, and 1.63% for polystyrene (PS), polybutylene succinate-co-adipate (PBSA), and polybutylene succinate (PBS), respectively. Trend analysis based on Henry's law illustrated that increasing pressure and decreasing temperature lead to an increase in carbon dioxide solubility. Finally, outlier discovery was applied using the leverage approach to detect the suspected data points. The outlier detection demonstrated the statistical validity of the developed correlations. William's plot of three generated correlations showed that all of the data points are located in the valid zone except one point for PBS polymer and three points for PS polymer.

In the recent years, application of different polymers has become an attractive issue in various industries including the petroleum industry. The fluid adsorption process in different polymers is a vital circumstance in the oil industry concepts such as enhanced oil recovery (EOR)^{1–3}, gas separation, imbibition of additives, and foaming processes^{4,5}. Carbon dioxide (CO₂) is one of the most significant gases, which plays a noteworthy role in polymers' structure, polymer foams, and production properties^{4,6}. Also, CO₂ and supercritical carbon dioxide (SCCO₂), (a supercritical carbon dioxide is described as a fluid for which both temperature and pressure are higher than critical values) have become one of the most conventional green materials, that have been extensively used in solvent, anti-solvent or a solute in numerous field processing including material synthesis, material modification, foaming processes, polymerization and particle production^{7–9}. SCCO₂ is potentially appealing as a solvent that shows properties that are a mixture of those commonly combined with liquids or gases. CO₂ solubility is the maximum CO₂ quantity that can solute in different solutions. Evaluation, prediction, and measurement of CO₂ solubility in different biodegradable polymers has become notable technology for engineers in various chemical applications such as extraction and generation of novel materials^{10–14}. Biodegradable polymers are a particular type of polymers that collapse by bacterial dissolution process to eventuate in natural fluids such as CO₂ and N₂. Poly butylene succinate (PBS) and polybutylene succinate-co-adipate (PBSA) are two applicable biodegradable polymers that have been generated by Showa Highpolymer Co. Ltd. and Showa Denko K.K.^{15,16}.

In order to predict gas solubilities in polymers, especially CO₂, various experimental, empirical, and theoretical approaches were investigated since 1986. In 1986 and 1993, Shah et al.^{17,18} measured solubility of different gases including CO₂ in silicone polymers at pressures up to 26 atmosphere and temperature values of 10, 35, and 55 °C. In 1994, Li et al.¹⁹ predicted the solubility of CO₂ in amine systems. They considered binary and ternary mixtures containing three solvents, namely mono-ethanolamine (MEA), methyl-diethanolamine (MDEA), and water (H₂O). They used temperature in a range of 0–225 °C. They modeled CO₂ solubility in amine mixtures as a function of temperature. Two years later, Sato et al.²⁰ investigated solubility of CO₂ and N₂ in polystyrene under high pressure and temperature conditions. They measured gas solubility at pressures up to 20 MPa and temperatures from 373.2 to 453.2 K. In 1998, Aubert²¹ calculated CO₂ solubility at pressures up to 9.65 MPa using

¹Department of Petroleum Engineering, Shahid Bahonar University of Kerman, Kerman, Iran. ²Département Etudes Thermodynamiques, Division Laboratoires, Sonatrach, Boumerdes, Algeria. ³Department of System Engineering, École de Technologie Supérieure, Montreal, QC, Canada. ⁴State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum (Beijing), Beijing, China. ✉email: mohammadhadi.shateri@etsmtl.ca; hemmati@uk.ac.ir

quartz crystal microbalance technique. Next year, Webb et al.²² and Sato et al.²³ evaluated diffusion and solubility of CO₂ in polymers under high pressures and temperatures. According to their research, the solubilities increased by increasing pressure and decreased by increasing temperature. In 2000, Sato et al.¹⁵ suggested empirical relations to determine solubility and diffusion coefficient of CO₂. They considered pressure and temperature as the dependent variables in the range of 1.025–20.144 MPa and 323.15–453.15 K, respectively. They achieved that solubility of CO₂ in molten state polymers increases by increasing pressure and decreasing temperature. A year later, Hilic et al.²⁴ measured solubility of N₂ and CO₂ in polystyrene, which considered pressure from 3.05 to 45 MPa and temperature from 338 to 402 K. In addition, an experimental technique with a vibrating-wire force sensor was applied. They got a linear relationship between increasing solubility with increasing pressure and decreasing temperature. In the same year, Sato et al.²⁵ calculated solubilities of CO₂ at the temperature range of 313.15–373.15 K and pressures up to 17.5 MPa. In 2002, Park et al.²⁶ studied about CO₂ solubility in alkanolamine solutions in the values of 40, 60 and 80 °C for temperature and 0.1–50 psia for pressure. They represented a vapor–liquid equilibrium of CO₂ in these solutions. In the same year, Sato et al.²⁷ examined CO₂ solubility in poly(2,6-dimethyl-1,4-phenylene ether) (PPO) and PS at temperatures of 373.15, 427.15, and 473.15 K and pressures up to 20 MPa. They obtained that solubility of CO₂ increases with increasing PPO concentration. A year later, in 2003, Hamed et al.²⁸ predicted the adsorption of CO₂ in various polymers based on a group contribution equation of state (EoS) with input ranges of 283–453 K and 1–200 bar for temperature and pressure, respectively. Their best result was an average absolute relative error (AARE) of 5.5% for polystyrene. In 2006, Li et al.²⁹ measured gas solubilities and diffusivities in polylactide at a temperature of 180–200 K and pressures up to 28 MPa using a magnetic suspension balance (MSB). Furthermore, they adopted a theoretical model based on Fick's second law to extract diffusion coefficients of N₂ and CO₂ in polylactide. They obtained that CO₂ exhibited lower diffusivity than N₂ at the same temperature. At that year, Nalawade et al.⁹ used SCCO₂ as a green solvent for processing polymer melts. They earned SCCO₂ is applicable in many polymerization processes due to its high solubility in polymers. In 2007, Lei et al.³⁰ generated buoyancy correlations and Sanchez and Lacombe equation of state to estimate CO₂ swelling degree, crystallinity, and solubility in polypropylene. They achieved CO₂ solubility first decreased and then increased with temperature. Two years later, Khajeh et al.³¹ developed intelligent model based on adaptive neuro fuzzy inference system (ANFIS) to predict solubility of CO₂ in polymers. They used up to 37 data points for different polymers. In 2011, Xu et al.³² investigated a theoretical study of solubility correlations of CO₂ in ether and carbonyl groups of polymers, namely poly(ethylene oxide) (PEO), poly(propylene oxide) (PPO), poly(vinyl acetate) (PVAc), poly(ethylene carbonate) (PEC) and poly(propylene carbonate) (PPC). They showed that the CO₂ solubility in PPC is higher than other polymers used in their study. Next year, Han et al.¹³ developed continuous reactions and considered economical concepts in SCCO₂ applications. In 2013, Li et al.³³ developed an artificial neural network (ANN) to estimate gas solubilities in polymers. Their research demonstrated good agreement between experimental and predicted data using their correlation. At the same year, Minelli and Sarti³⁴ measured solubility and permeability of CO₂ in various glassy polymers by considering diffusion coefficient as a kinetic factor. In 2015, different mathematical and theoretical approaches by Ting and Yuan¹⁰, Li et al.⁷ and Quan et al.¹² were studied to estimate CO₂ properties including solubility. All of them showed that the CO₂ solubility has direct relation with pressure and reverse relation with temperature. Two years later, Mengshan et al.^{8,35} developed an artificial neural network and artificial intelligence technique based on diffusion theory to predict solubility of CO₂ and SCCO₂ in polymers. In 2019, Soleimani et al.⁴ developed decision tree (DT) based smart model for estimating solubility of CO₂. They used 515 data points with a range of 306–483.7 K for temperature and 1.025–44.41 MPa for pressure. One year later, Li et al.³⁶ investigated a comprehensive review of CO₂ polymer system. They used two types of multi-scaled methods, namely thermodynamic-calculation model and computer simulation to measure CO₂ solubility in polymers. Their developed model can be utilized in chemistry and chemical industries, such as phase rheological property and polymer self-assembly. In 2022, various experimental, theoretical, and modeling researches have been done in order to measure solubility of CO₂ and other gases in water-polymer systems. Sun et al.³⁷ measured CO₂ solubility in oil-based and water-based drilling fluids using the sample analysis approach. Their results indicated that the salting-out effect of electrolyte on gas solubility can be increased with increasing the molar concentration of ions. Their study also showed that the errors of CO₂ solubility in the oil-based and water-based drilling fluids are 6.75% and 3.47%, respectively. Besides, Ushiki et al.³⁸ evaluated CO₂ solubility and diffusivity in polycaprolactone (PCL) performing perturbed-chain statistical associating fluid theory (PC-SAFT) and free volume methods. According to their work, CO₂ solubility was recognized to conform with Henry's law, and the PC-SAFT EoS sufficiently described the solubility. Also, Kiran et al.³⁹ assessed diffusivity and solubility of CO₂ and N₂ in polymers. They used Sanchez-Lacombe EoS in modeling solubility. Furthermore, Ricci et al.⁴⁰ provided a comprehensive theoretical framework for the supercritical sorption and transport of CO₂ in polymers. In their study, CO₂ sorption was modelled utilizing data available across the critical region, at different temperatures and pressures up to 18 MPa.

The present research mostly focuses on generating accurate correlations for CO₂ solubility prediction considering the pressure and temperature of the polymer as input variables. The generated correlations are based on gene expression programming (GEP) technique. A comprehensive databank including of 53 data points for PBS, 43 data points for PBSA and 92 data points for PS polymer is collected^{15,20,24,25}. After generating correlations, statistical and graphical error tests are applied to assess the accuracy of the correlations. Likewise, the capability of the represented correlations in predicting the real trend of the CO₂ solubility with the change of pressure and temperature is appraised. Lately, the leverage approach is performed to detect the outlier data points in the dataset.

Data collection

In this research, GEP algorithm was implemented to predict the amount of CO₂ solubility in three different polymers, namely PBS, PBSA, and polystyrene (PS). For this aim, 53 data points for PBS, 43 data points for PBSA, and 92 data points for PS polymer were collected^{15,20,24,25}. In this work, pressure and temperature of carbon dioxide were considered as input parameters. A summary of the gathered data points is shown in Table 1. As pointed up in Table 1, extensive ranges of temperature and pressure of CO₂ are supplied in this study.

Correlation development

In order to generate CO₂ solubility correlations, Gene expression programming (GEP) evolutionary algorithm has been applied. GEP which was firstly proposed by Ferreira in 2001⁴¹, is a normally comprehensive phenotype technique in which the chromosomes form a correctly inseparable, operative entity⁴². This technique is extensively used in computer programming and modeling applications^{43–46}. Gene expression programming algorithms are complicated tree-based structures that coordinate by changing their shape, composition and sizes. By encoding trees as vectors of symbols and transforming them into them just in order to assess their fitness, this technique can indirectly produce trees⁴⁷. This soft computing technique is strong predictive algorithm that is widely used for various field application purposes. Commonly, the GEP technique has two components, namely chromosome and the expression trees (ETs). The possible solutions are encoded by the chromosomes and is regarded as the linear string with particular length, hence these solutions will be decoded into the real candidate solution termed expression tree⁴⁸. After producing of chromosomes of first-production individuals and choosing them based on fitness function to re-generate with modifications, new generation individuals were presented to the developmental operation of selection environment confrontation, genome expression, and modified reproduce⁴⁹. Additionally, gene expression programming automatically creates algebraic expressions to answer nonlinear problems⁵⁰. The schematic flowchart of GEP procedure is depicted in Fig. 1.

Results and discussion

Development of correlations. In the present study, gene expression programming tree-based soft computing approach was carried out to develop accurate correlations for predicting CO₂ solubility in different polymers. The developed correlations consider CO₂ solubility as a function of pressure and temperature of cor-

Polymer	Pressure (MPa)	Temperature (K)	CO ₂ solubility (g/g)	Number of data
PBS	1.025–20.144	323.15–453.15	0.00876–0.1761	53
PBSA	1.098–20.074	323.15–453.15	0.01184–0.1741	43
PS	2.068–44.41	338.22–473.15	0.00714–0.16056	92

Table 1. Summary of experimental data points utilized in this work.

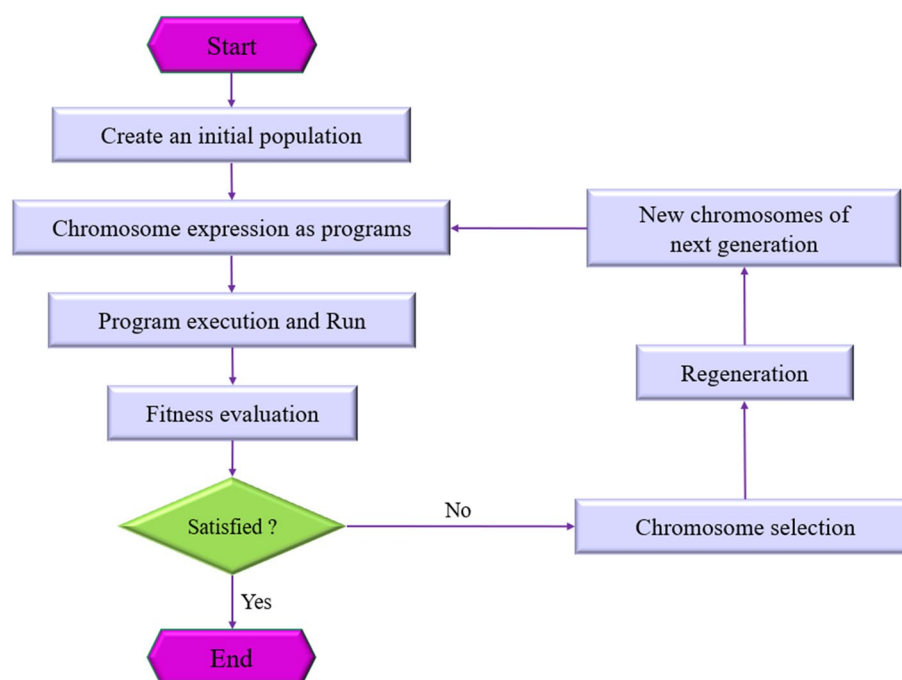


Figure 1. The schematic framework of the gene expression programming (GEP).

responding polymer and use them as input variables. To generate accurate and user-friendly correlations, an exhaustive databank consists of 53 data points for PBS polymer, 43 data points for PBSA polymer and 92 data points for PS polymer was collected from previous literature. Table 2 represents the GEP parameters utilized in this research.

Using the aforementioned approach, the final formulas for the determination of CO₂ solubility based on gene expression programming technique, are listed below:

CO₂ solubility correlation in PBS polymer.

$$S_{CO_2} = \frac{2.403631 \times P \times \left(T + (P + 2 \times T) \times \cos \left(\cos \left(\frac{T}{\cos \left(\cos \left(\frac{8}{T} \right)} \right) \right) \right) \right)}{T \times (P + 2 \times T)} - 0.000401 \quad (1)$$

CO₂ solubility correlation in PBSA polymer.

$$S_{CO_2} = 0.009492 + \frac{59525.424282 \times ((P - 8) \times (P + T) \times \text{sigmoid}(P) + 16 \times (P - \text{sigmoid}(P)) \times (T + 0.83907) \times \cos(\text{sigmoid}(P)))}{T^2 \times (P + T) \times (T + 0.83907) \times \text{sigmoid}(P)} \quad (2)$$

CO₂ solubility correlation in PS polymer.

$$S_{CO_2} = \frac{980.876013 \times P}{(T - \text{sigmoid}(P)) \times (8 \times P + T - 72 \times \text{sigmoid}(P) - 48)} + \frac{13.623278 \times \cos(\cos(T))}{T \times \text{sigmoid}(P)} - 0.040355 \quad (3)$$

where P and T denote pressure and temperature of aforementioned polymers, respectively. In the above correlations, the units of P and T are MPa and K, respectively. The generated correlations in this study are applicable for CO₂ solubility prediction in various ranges of temperature and pressure of the mentioned polymers.

Statistical performance assessment. In order to show and compare the precision of the generated correlations, some important statistical parameters including root mean square error (RMSE), standard deviation (SD), coefficient of determination (R²), the average relative error (ARE) and the average absolute relative error (AARE) were applied⁵¹. These terms are given below:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (S(\text{exp}) - S(\text{cal}))^2} \quad (4)$$

$$SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n \left(\frac{S(\text{exp}) - S(\text{cal})}{S(\text{exp})} \right)^2} \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (S(\text{exp}) - S(\text{cal}))^2}{\sum_{i=1}^n (S(\text{cal}) - S(\text{avg}))^2} \quad (6)$$

$$ARE = \frac{1}{n} \sum_{i=1}^n (E_i) \quad (7)$$

where E_i is the partial deviation that is described as:

Parameters	Value/setting
The number of head size	10–15
Chromosome	500
Number of generation	300
Mutation rate	0.25
Inversion rate	0.1
Operators used	+, −, ×, /, exp, X ² , INV, cos, ln, sqrt

Table 2. GEP setting parameters used in the study.

$$E_i = \left[\frac{S(\text{exp}) - S(\text{cal})}{S(\text{exp})} \right] \times 100 \quad (8)$$

$$\text{AARE} = \frac{1}{n} \sum_{i=1}^n |E_i| \quad (9)$$

where n , $S(\text{exp})$, $S(\text{cal})$ and $S(\text{avg})$ are the number of data, actual CO₂ solubility value, calculated CO₂ solubility value, and the average of the actual data points, respectively. The prementioned statistical parameters for the three generated correlations are detailed for the training, testing, and whole datasets in Table 3. As described in this table, the AARE of the correlation for the PBS polymer is lower than other two correlations generated in this work. Results demonstrate that generated correlation for the PBS polymer has the lowest standard deviation (0.028) and RMSE (0.00178). However, the correlations developed for the other two polymers also have acceptable accuracy. As presented in Table 3, the AARE values for PBS and PBSA polymers were obtained less than AARE for PS polymer, which was due to the nature of the experimental data related to PS polymer. It is obvious that the generated correlations are reliable and sometimes, due to the nature of the experimental data values of different materials (like polymers), different error values may be obtained.

Graphical performance assessment. This section represents a graphical description of the comparison among the results of the generated correlations and the actual data. The predicted CO₂ solubility values in PBS polymer are sketched versus actual ones in Fig. 2a. Likewise, the predicted CO₂ solubility values in PBSA and PS polymers are depicted versus experimental data in Fig. 2b,c, respectively. The closer the sketched data points to the 45° line, the greater the uniformity of the correlations is. According to these plots, it is apparent that the results of the generated user-friendly correlations illustrate satisfactory agreement around the ideal line. Additionally, the relative error curves of the developed correlations of the CO₂ solubility in PBS polymer, PBSA polymer, and PS polymer are presented in Fig. 3a–c, respectively.

Furthermore, to show the accuracy of the presented correlations in different ranges of pressure and temperature, the correlations' performances in terms of AARE were sketched against five sets of pressure and three sets of temperature. Figure 4 demonstrates the AARE of the correlations in different ranges of input parameters. For various ranges of pressure, the correlation of CO₂ solubility in PBS polymer clarifies a steady performance and its AARE is lower than 2.9% in all ranges. Besides, a reliable performance can be perceived from the correlation of CO₂ solubility in PBS polymer up to the last temperature range. This figure validates the efficiency of the developed correlation of CO₂ solubility in PBS polymer over other developed correlations in the present study.

Afterwards, the cumulative frequency analysis of the absolute percent relative error (APRE) for the generated correlations in this work is shown in Fig. 5. According to the results of this figure, the correlation of CO₂ solubility in PBS polymer could estimate more than 90% of CO₂ solubility values with an APRE of less than 5%, and also more than 98% of the CO₂ solubility values by the correlation for PBS polymer have an AARE of less than 10%.

Additionally, absolute relative error comparison among generated correlations was carried out. Figure 6 describes the AARE comparison between the prementioned correlations. According to this figure, the developed correlation of CO₂ solubility in PBS polymer revealed the highest accuracy and the lowest AARE between other correlations generated in this research.

Trend analysis of the generated correlations. Trend analysis is a well-known applicable technique to visualize the output variation with the change of input variables^{52,53}. The predictions of the CO₂ solubility correlations are depicted versus temperature and pressure in Fig. 7 to investigate the capability of the generated correlations in following the actual expected trends of CO₂ solubility values with the change of pressure and temperature. According to Henry's law, it is evident that CO₂ solubility increases with decreasing temperature and increasing pressure⁵⁴. Carbon dioxide has a propensity, namely plasticizing effect⁵⁵. It means that the molecules of CO₂ are pressured in the chains of the polymer as a consequence of increasing pressure, which results in an extension of the pore space within the molecules and, then, for this reason, an addition of their movement^{56,57}.

Polymer	Status	ARE (%)	AARE (%)	RMSE	SD	R ²
PBS	Train	0.217	1.775	0.001	0.031	0.997
	Test	0.510	1.097	0.000	0.014	0.999
	All	0.278	1.635	0.001	0.028	0.998
PBSA	Train	1.660	1.271	0.002	0.092	0.995
	Test	– 4.402	8.175	0.004	0.099	0.989
	All	0.391	5.879	0.003	0.093	0.993
PS	Train	3.305	9.197	0.004	0.156	0.991
	Test	4.274	11.819	0.005	0.242	0.973
	All	3.494	9.710	0.004	0.176	0.989

Table 3. Statistical assessment of the generated correlations in this work.

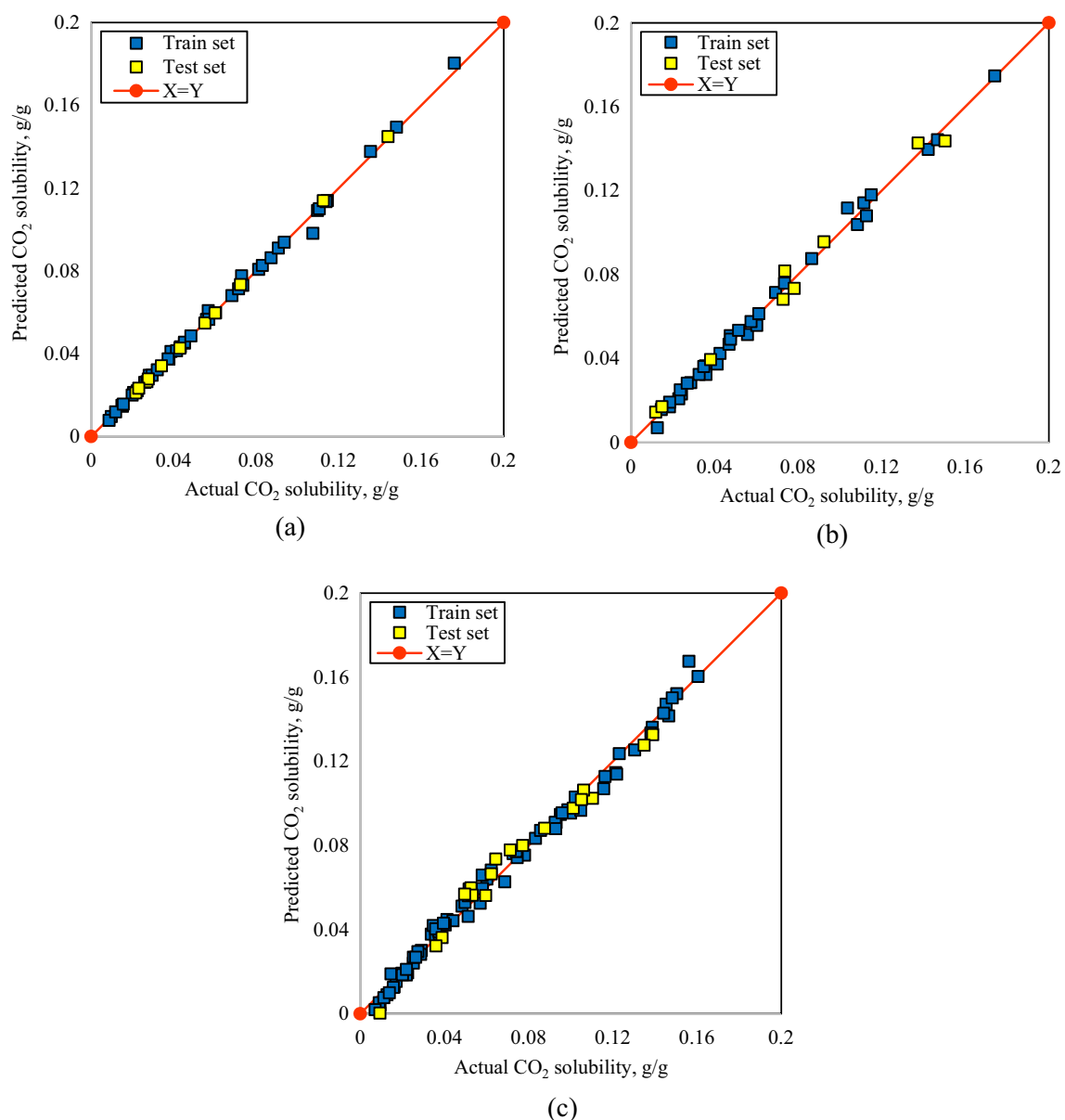


Figure 2. Cross plots of the predicted and experimental CO₂ solubility values in (a) PBS, (b) PBSA, (c) PS polymers.

This causes it feasible to absorb more gas molecules. Likewise, by decreasing the temperature the CO₂ molecules obtain lower kinetic energy and they do not have a tendency for releasing from the solution and for staying in a condition with more independence⁵⁸. As a consequence, the solubility would increase.

Outlier discovery of the developed correlations. Outlier discovery plays an important role to identify data that may vary from other data points exist in a dataset⁵⁹. The leverage technique is a trustworthy method for outlier discovery which concerns with the values of the standardized residuals and a matrix, namely the Hat matrix made of the actual and the predicted values obtained from the correlations⁶⁰. According to this approach, if most of the data points located in the ranges of $-3 \leq R \leq 3$ (R denotes the standardized residual) and $0 \leq H_i \leq H^*$, it illustrates that the results of the generated correlations are dependable and valid^{61–63}. Figures 8, 9 and 10 represent William plots of the generated correlations of CO₂ solubility in PBS, PBSA, and PS polymers, respectively. For PBS polymer it is obvious that all of the data points placed in a valid zone except one. Also, the results of the generated correlation of PBSA polymer show that all of the data points located in a valid region. At the end, Fig. 10 presents a William plot of CO₂ solubility correlation in PS polymer, showing that among whole dataset consists of 92 data points used for this polymer, only 3 data points are recognized as out of leverage data points.

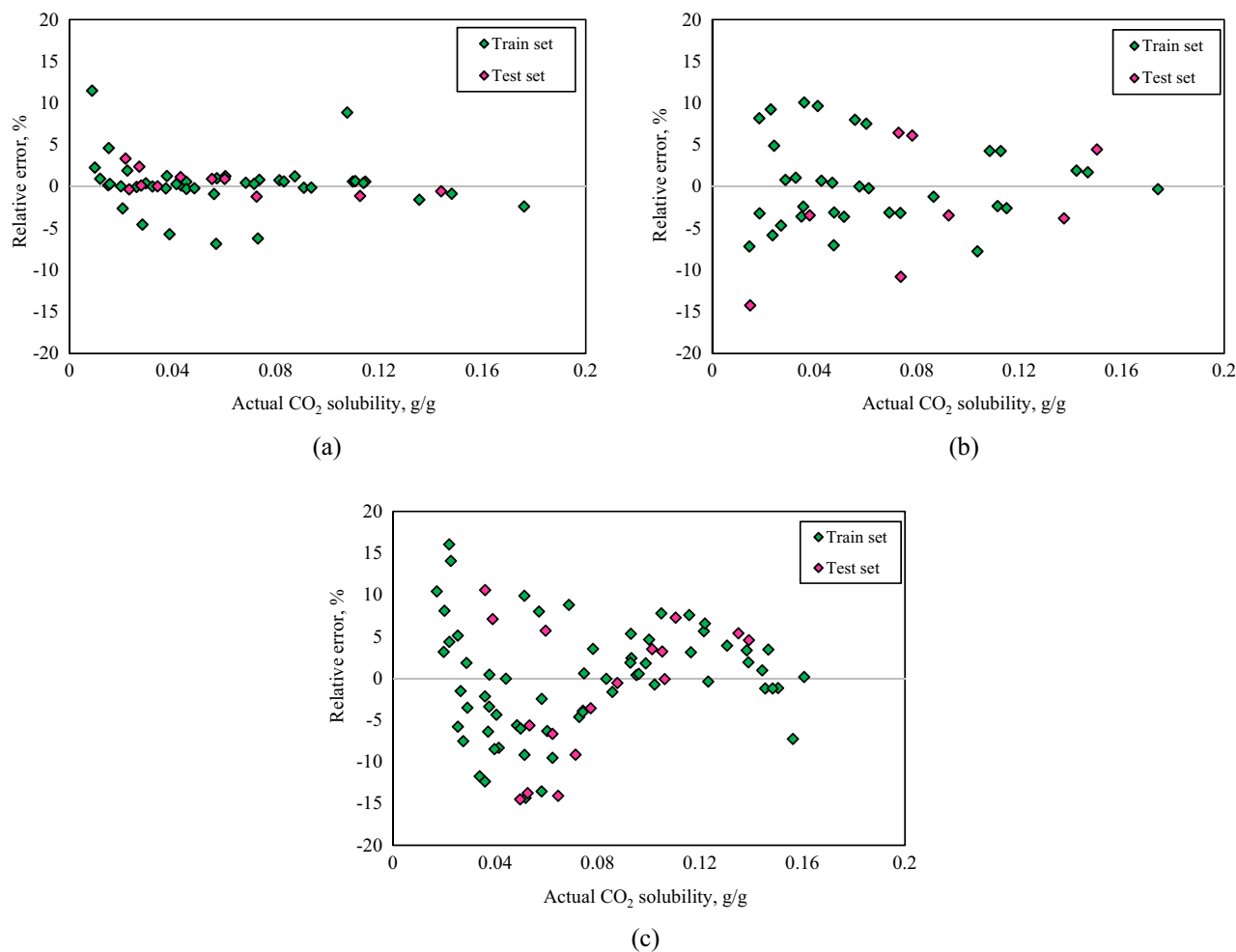


Figure 3. Relative error distribution curves of the generated correlation of the CO₂ solubility in (a) PBS, (b) PBSA, (c) PS polymers.

Conclusions

The present research aimed to predict CO₂ solubility as a strong effective parameter in polymerization processes. PBS, PBSA, and PS were three polymers, which were utilized in this work. For this purpose, gene expression programming (GEP) technique was applied. To this aim, a widespread dataset was gathered from previous literature. Results showed that the generated correlation of CO₂ solubility for PBS polymer could present the highest accuracy in predicting solubility of CO₂ with an AARE of 1.63%, SD of 0.028, and RMSE of 0.001. The sketched CO₂ solubility curves using the trend analysis demonstrated that all three generated correlations in this study could exactly fit the actual trends of CO₂ solubility variation. The simple generated correlations can be performed in wide ranges of pressures and temperatures and represent high accuracy. The leverage approach showed that all the data points seem to be reliable and valid except four, which were placed in a lower suspected and out of leverage zones. In order to precisely simulate CO₂ solubility in polymers in a future works, it is recommended to generate new correlations, and also develop intelligent schemes.

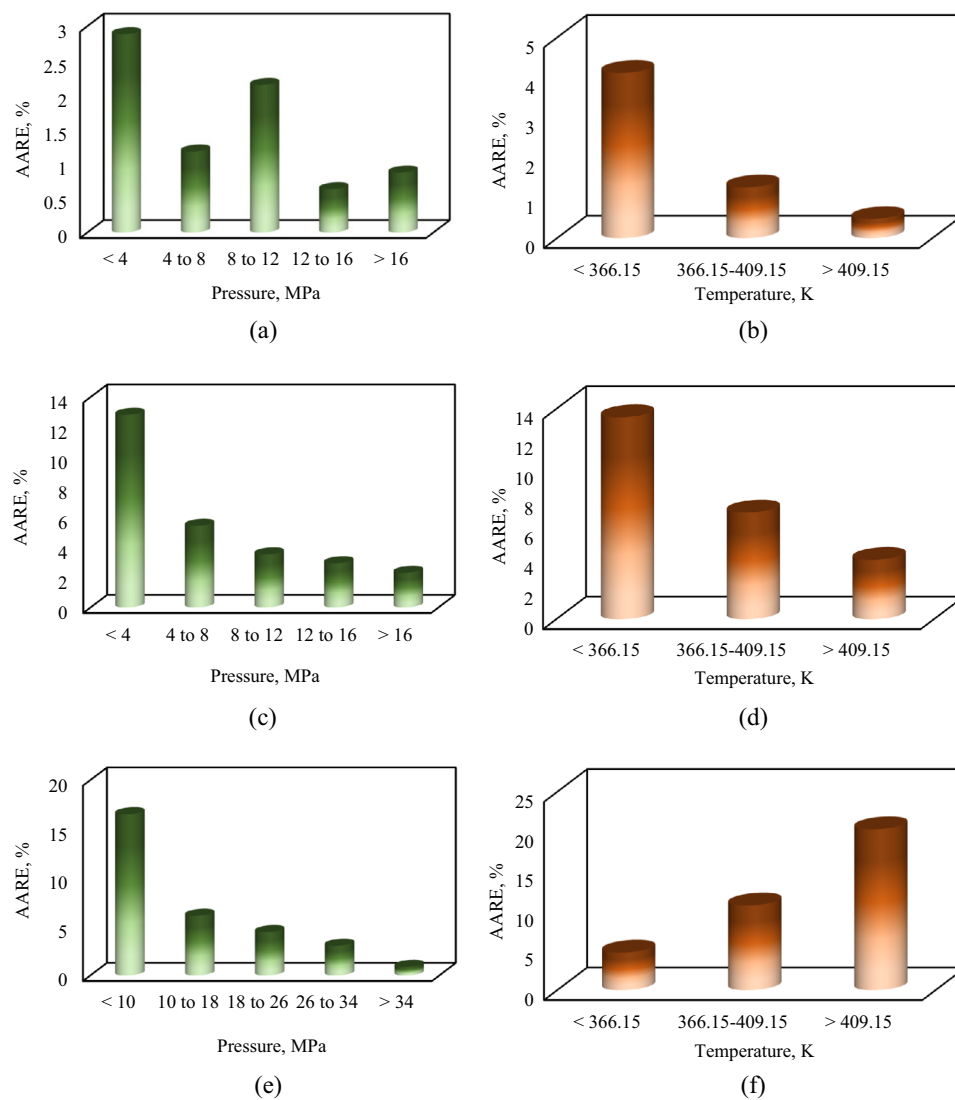


Figure 4. AARE for the different correlations performed in this research for the three polymers in various inputs ranges. (a,b) PBS; (c,d) PBSA; (e,f) PS.

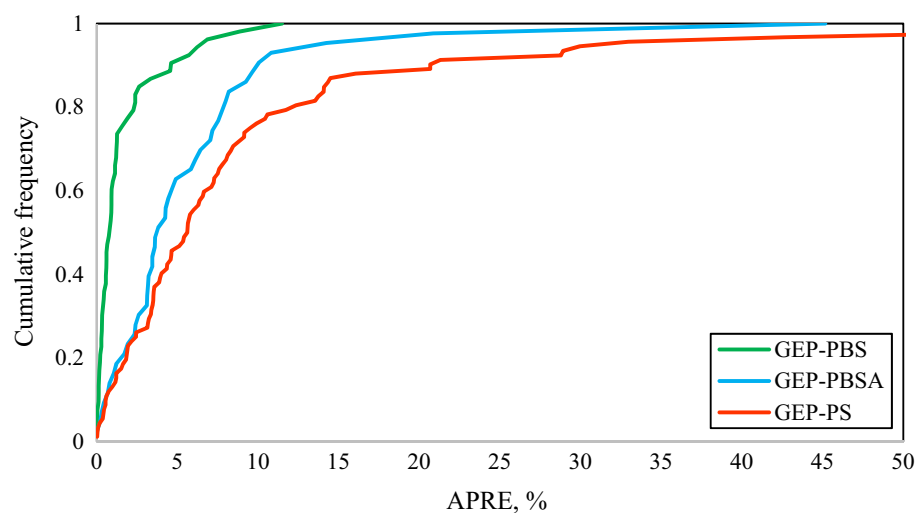


Figure 5. Cumulative frequency plot of generated correlations in this study.

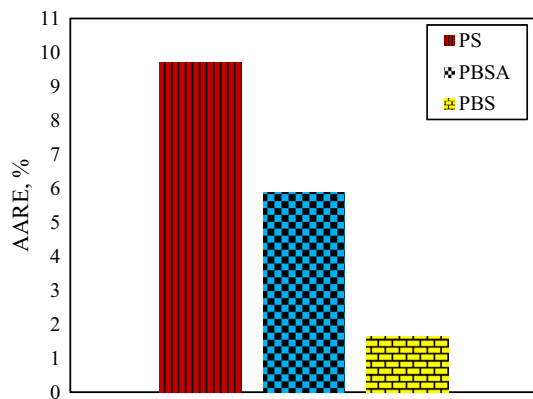
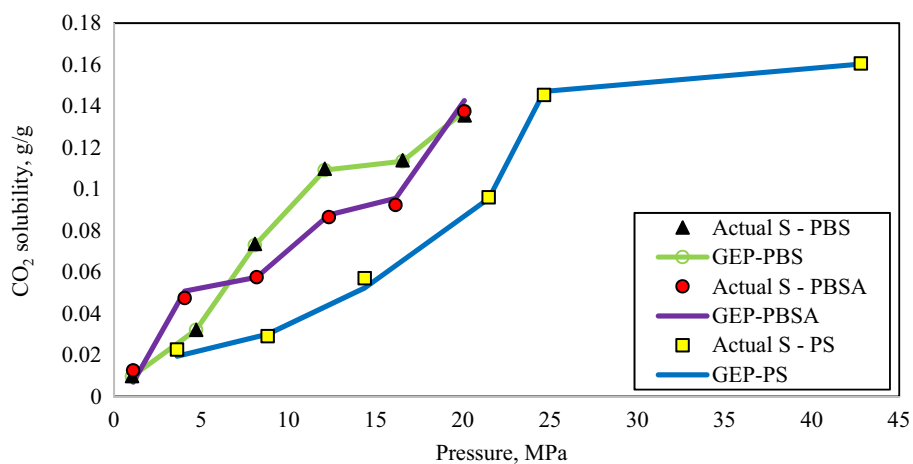
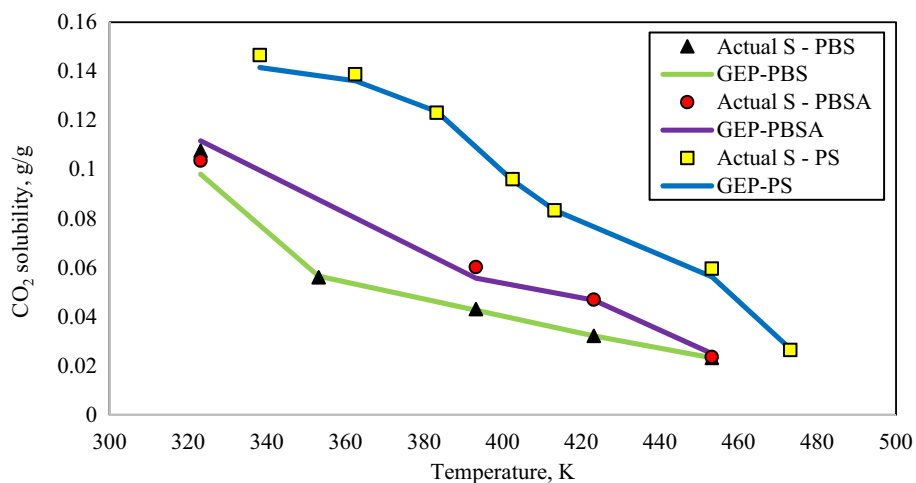


Figure 6. Comparison among AARE values of the implemented correlations.



(a)



(b)

Figure 7. Comparison of the CO₂ solubility variation for the generated correlations in this work with actual data. (a) CO₂ solubility change with pressure; (b) CO₂ solubility change with temperature.

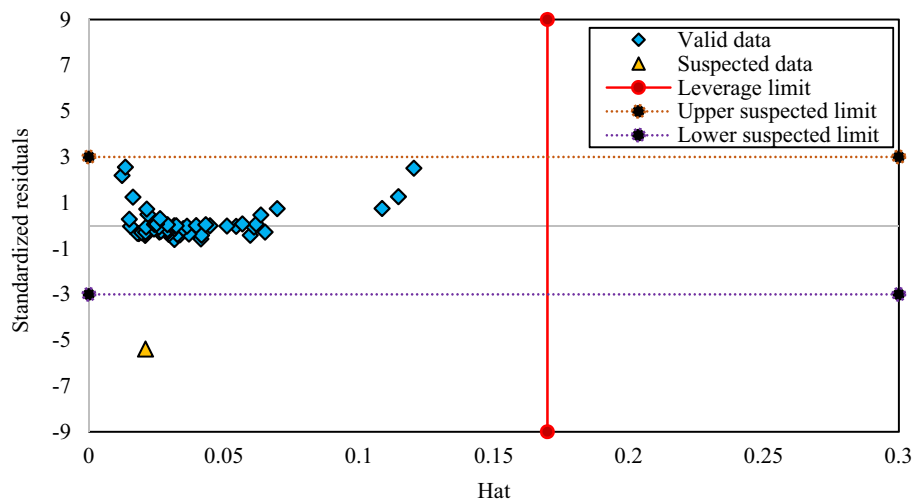


Figure 8. The William plot of the generated correlation for PBS polymer.

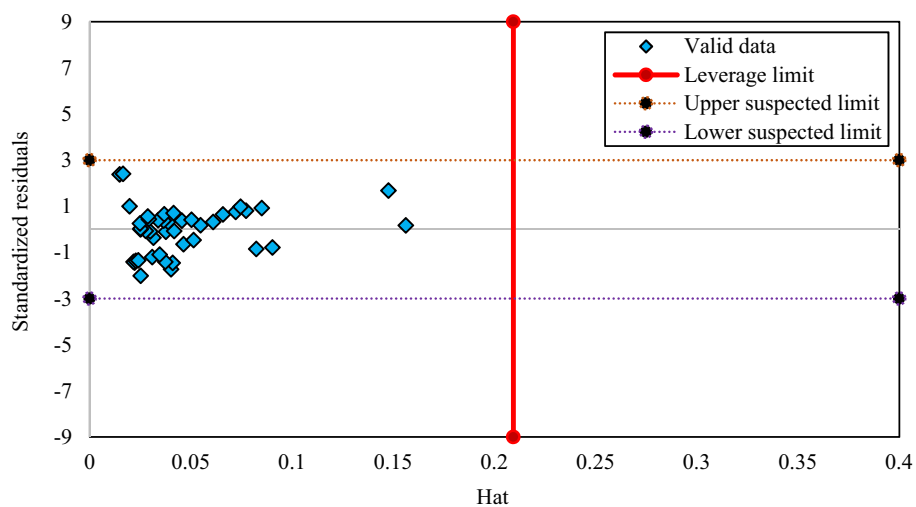


Figure 9. The William plot of the generated correlation for PBSA polymer.

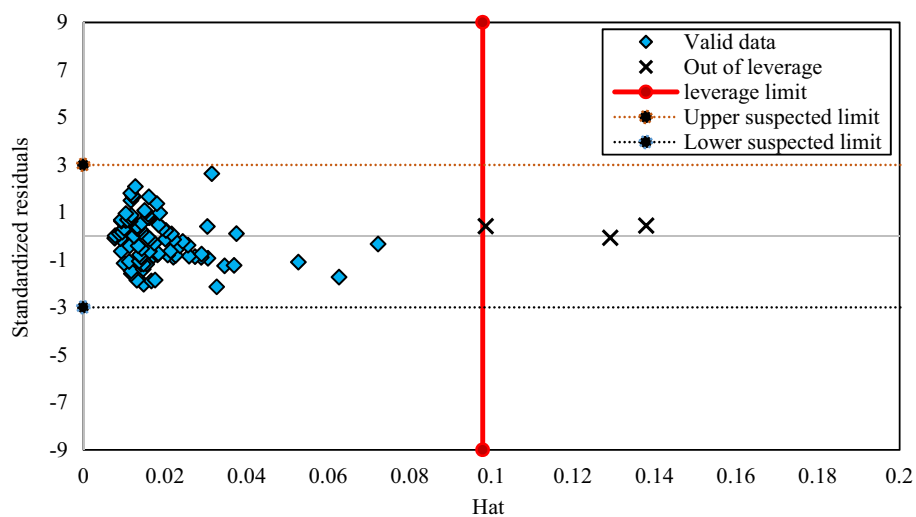


Figure 10. The William plot of the generated correlation for PS polymer.

Data availability

All the data are available from the corresponding author on reasonable request.

Received: 11 February 2023; Accepted: 24 July 2023

Published online: 02 August 2023

References

- Sheng, J. J. *Enhanced Oil Recovery Field Case Studies* (Gulf Professional Publishing, 2013).
- Thomas, S. Enhanced oil recovery—an overview. *Oil Gas Sci. Technol. Rev. l'IFP* **63**, 9–19 (2008).
- Divandari, H., Amiri-Ramsheh, B. & Zabihi, R. Steam flooding (steam drive). *Thermal Methods* **20**, 47 (2023).
- Soleimani, R. *et al.* Evolving an accurate decision tree-based model for predicting carbon dioxide solubility in polymers. *Chem. Eng. Technol.* **43**, 514–522 (2020).
- Li, D.-C., Liu, T., Zhao, L. & Yuan, W.-K. Solubility and diffusivity of carbon dioxide in solid-state isotactic polypropylene by the pressure-decay method. *Ind. Eng. Chem. Res.* **48**, 7117–7124 (2009).
- Zheng, H., Mahmoudzadeh, A., Amiri-Ramsheh, B. & Hemmati-Sarapardeh, A. Modeling viscosity of CO₂-N₂ gaseous mixtures using robust tree-based techniques: Extra tree, random forest, GBoost, and LightGBM. *ACS Omega* **8**, 13863–13875 (2023).
- Li, M. *et al.* Solubility prediction of supercritical carbon dioxide in 10 polymers using radial basis function artificial neural network based on chaotic self-adaptive particle swarm optimization and K-harmonic means. *RSC Adv.* **5**, 45520–45527 (2015).
- Mengshan, L. *et al.* Prediction of supercritical carbon dioxide solubility in polymers based on hybrid artificial intelligence method integrated with the diffusion theory. *RSC Adv.* **7**, 49817–49827 (2017).
- Nalawade, S. P., Picchioni, F. & Janssen, L. Supercritical carbon dioxide as a green solvent for processing polymer melts: Processing aspects and applications. *Prog. Polym. Sci.* **31**, 19–43 (2006).
- Ru-Ting, X. & Xing-Yuan, H. Predictive calculation of carbon dioxide solubility in polymers. *RSC Adv.* **5**, 76979–76986 (2015).
- Zhang, Q., Vanparijs, N., Louage, B., De Geest, B. G. & Hoogenboom, R. Dual pH- and temperature-responsive RAFT-based block co-polymer micelles and polymer-protein conjugates with transient solubility. *Polym. Chem.* **5**, 1140–1144 (2014).
- Quan, S. *et al.* A bio-inspired CO₂-philic network membrane for enhanced sustainable gas separation. *J. Mater. Chem. A* **3**, 13758–13766 (2015).
- Han, X. & Poliakov, M. Continuous reactions in supercritical carbon dioxide: Problems, solutions and possible ways forward. *Chem. Soc. Rev.* **41**, 1428–1436 (2012).
- Chandra, R. & Rustgi, R. Biodegradable polymers. *Prog. Polym. Sci.* **23**, 1273–1335 (1998).
- Sato, Y. *et al.* Solubility and diffusion coefficient of carbon dioxide in biodegradable polymers. *Ind. Eng. Chem. Res.* **39**, 4813–4819 (2000).
- Nishioka, M., Tuzuki, T., Wanajyo, Y., Oonami, H. & Horiuchi, T. *Studies in Polymer Science* Vol. 12, 584–590 (Elsevier, 1994).
- Yampolskii, Y. & Paterson, R. Solubility of gases in polymers. *Exp. Determin. Solubil.* **6**, 151–171 (2003).
- Shah, V., Hardy, B. & Stern, S. Solubility of carbon dioxide, methane, and propane in silicone polymers. Effect of polymer backbone chains. *J. Polym. Sci. Part B Polym. Phys.* **31**, 313–317 (1993).
- Li, Y.-G. & Mather, A. E. Correlation and prediction of the solubility of carbon dioxide in a mixed alkanolamine solution. *Ind. Eng. Chem. Res.* **33**, 2006–2015 (1994).
- Sato, Y., Yurugi, M., Fujiwara, K., Takishima, S. & Masuoka, H. Solubilities of carbon dioxide and nitrogen in polystyrene under high temperature and pressure. *Fluid Phase Equilib.* **125**, 129–138 (1996).
- Aubert, J. H. Solubility of carbon dioxide in polymers by the quartz crystal microbalance technique. *J. Supercrit. Fluids* **11**, 163–172 (1998).
- Webb, K. F. & Teja, A. S. Solubility and diffusion of carbon dioxide in polymers. *Fluid Phase Equilib.* **158**, 1029–1034 (1999).
- Sato, Y., Fujiwara, K., Takikawa, T., Takishima, S. & Masuoka, H. Solubilities and diffusion coefficients of carbon dioxide and nitrogen in polypropylene, high-density polyethylene, and polystyrene under high pressures and temperatures. *Fluid Phase Equilib.* **162**, 261–276 (1999).
- Hilic, S., Boyer, S. A., Pádua, A. A. & Grolier, J. P. E. Simultaneous measurement of the solubility of nitrogen and carbon dioxide in polystyrene and of the associated polymer swelling. *J. Polym. Sci. Part B Polym. Phys.* **39**, 2063–2070 (2001).
- Sato, Y., Takikawa, T., Takishima, S. & Masuoka, H. Solubilities and diffusion coefficients of carbon dioxide in poly (vinyl acetate) and polystyrene. *J. Supercrit. Fluids* **19**, 187–198 (2001).
- Park, S. H., Lee, K. B., Hyun, J. C. & Kim, S. H. Correlation and prediction of the solubility of carbon dioxide in aqueous alkanolamine and mixed alkanolamine solutions. *Ind. Eng. Chem. Res.* **41**, 1658–1665 (2002).
- Sato, Y., Takikawa, T., Yamane, M., Takishima, S. & Masuoka, H. Solubility of carbon dioxide in PPO and PPO/PS blends. *Fluid Phase Equilib.* **194**, 847–858 (2002).
- Hamed, M., Muralidharan, V., Lee, B. & Danner, R. P. Prediction of carbon dioxide solubility in polymers based on a group-contribution equation of state. *Fluid Phase Equilib.* **204**, 41–53 (2003).
- Li, G., Li, H., Turng, L., Gong, S. & Zhang, C. Measurement of gas solubility and diffusivity in polylactide. *Fluid Phase Equilib.* **246**, 158–166 (2006).
- Lei, Z., Ohyabu, H., Sato, Y., Inomata, H. & Smith, R. L. Jr. Solubility, swelling degree and crystallinity of carbon dioxide-polypropylene system. *J. Supercrit. Fluids* **40**, 452–461 (2007).
- Khajeh, A., Modarress, H. & Rezaee, B. Application of adaptive neuro-fuzzy inference system for solubility prediction of carbon dioxide in polymers. *Expert Syst. Appl.* **36**, 5728–5732 (2009).
- Xu, M., Chen, J., Zhang, C., Du, Z. & Mi, J. A theoretical study of structure-solubility correlations of carbon dioxide in polymers containing ether and carbonyl groups. *Phys. Chem. Chem. Phys.* **13**, 21084–21092 (2011).
- Li, M. *et al.* Prediction of gas solubility in polymers by back propagation artificial neural network based on self-adaptive particle swarm optimization algorithm and chaos theory. *Fluid Phase Equilib.* **356**, 11–17 (2013).
- Minelli, M. & Sarti, G. C. Permeability and solubility of carbon dioxide in different glassy polymer systems with and without plasticization. *J. Membr. Sci.* **444**, 429–439 (2013).
- Mengshan, L., Wei, W., Bingsheng, C., Yan, W. & Xingyuan, H. Solubility prediction of gases in polymers based on an artificial neural network: A review. *RSC Adv.* **7**, 35274–35282 (2017).
- Li, M. *et al.* Models for the solubility calculation of a CO₂/polymer system: A review. *Mater. Today Commun.* **25**, 101277 (2020).
- Sun, X. *et al.* Experiments and modeling of CO₂ solubility in water-based and oil-based drilling fluids. *J. Petrol. Sci. Eng.* **212**, 110336 (2022).
- Ushiki, I., Kawashima, H., Kihara, S.-I. & Takishima, S. Solubility and diffusivity of supercritical CO₂ for polycaprolactone in its molten state: Measurement and modeling using PC-SAFT and free volume theory. *J. Supercrit. Fluids* **181**, 105499 (2022).
- Kiran, E., Sarver, J. A. & Hassler, J. C. Solubility and diffusivity of CO₂ and N₂ in polymers and polymer swelling, glass transition, melting, and crystallization at high pressure: A critical review and perspectives on experimental methods, data, and modeling. *J. Supercrit. Fluids* **185**, 105378 (2022).

40. Ricci, E., De Angelis, M. G. & Minelli, M. A comprehensive theoretical framework for the sub and supercritical sorption and transport of CO₂ in polymers. *Chem. Eng. J.* **435**, 135013 (2022).
41. Ferreira, C. Gene expression programming: A new adaptive algorithm for solving problems. *arXiv preprint cs/0102027* (2001).
42. Umar, A. A., Saaid, I. M., Sulaimon, A. A. & Pilus, R. M. Predicting the viscosity of petroleum emulsions using gene expression programming (GEP) and response surface methodology (RSM). *J. Appl. Math.* **20**, 20 (2020).
43. Zhong, J., Feng, L. & Ong, Y.-S. Gene expression programming: A survey. *IEEE Comput. Intell. Mag.* **12**, 54–72 (2017).
44. Amar, M. N., Larestani, A., Lv, Q., Zhou, T. & Hemmati-Sarapardeh, A. Modeling of methane adsorption capacity in shale gas formations using white-box supervised machine learning techniques. *J. Petrol. Sci. Eng.* **208**, 109226 (2022).
45. Amar, M. N. Prediction of hydrate formation temperature using gene expression programming. *J. Nat. Gas Sci. Eng.* **89**, 103879 (2021).
46. Amar, M. N., Ghriga, M. A., Seghier, M. E. A. B. & Ouair, H. Predicting solubility of nitrous oxide in ionic liquids using machine learning techniques and gene expression programming. *J. Taiwan Inst. Chem. Eng.* **128**, 156–168 (2021).
47. Baniasadi, H., Kamari, A., Heidararabi, S., Mohammadi, A. H. & Hemmati-Sarapardeh, A. Rapid method for the determination of solution gas-oil ratios of petroleum reservoir fluids. *J. Nat. Gas Sci. Eng.* **24**, 500–509 (2015).
48. Rostami, A., Arabloo, M., Kamari, A. & Mohammadi, A. H. Modeling of CO₂ solubility in crude oil during carbon dioxide enhanced oil recovery using gene expression programming. *Fuel* **210**, 768–782 (2017).
49. Mirzaie, M. & Tatar, A. Modeling of interfacial tension in binary mixtures of CH₄, CO₂, and N₂-alkanes using gene expression programming and equation of state. *J. Mol. Liq.* **320**, 114454 (2020).
50. Traore, S., Luo, Y. & Fipps, G. Gene-expression programming for short-term forecasting of daily reference evapotranspiration using public weather forecast information. *Water Resour. Manage* **31**, 4891–4908 (2017).
51. Sarapardeh, A. H., Larestani, A., Menad, N. A. & Hajirezaie, S. *Applications of Artificial Intelligence Techniques in the Petroleum Industry* (Gulf Professional Publishing, 2020).
52. Kirk, H., Haynes, F. & Monroe, R. Application of trend analysis to horticultural field trials. *J. Am. Soc. Hortic. Sci.* **20**, 20 (1980).
53. Amiri-Ramsheh, B., Safaei-Farouji, M., Larestani, A., Zabihi, R. & Hemmati-Sarapardeh, A. Modeling of wax disappearance temperature (WDT) using soft computing approaches: Tree-based models and hybrid models. *J. Petrol. Sci. Eng.* **208**, 109774 (2022).
54. Kumeľan, J., Kamps, A.P.-S., Tuma, D. & Maurer, G. Solubility of CO₂ in the ionic liquid [hmim][Tf₂N]. *J. Chem. Thermodyn.* **38**, 1396–1401 (2006).
55. Minelli, M., Oradei, S., Fiorini, M. & Sarti, G. C. CO₂ plasticization effect on glassy polymeric membranes. *Polymer* **163**, 29–35 (2019).
56. Khoshraftar, Z. & Ghaemi, A. Prediction of CO₂ solubility in water at high pressure and temperature via deep learning and response surface methodology. *Case Stud. Chem. Environ. Eng.* **7**, 100338 (2023).
57. Messabeh, H., Contamine, F., Cézac, P., Serin, J. P. & Gaucher, E. C. Experimental measurement of CO₂ solubility in aqueous NaCl solution at temperature from 323.15 to 423.15 K and pressure of up to 20 MPa. *J. Chem. Eng. Data* **61**, 3573–3584 (2016).
58. Thibault, Y. & Holloway, J. R. Solubility of CO₂ in a Ca-rich leucite: Effects of pressure, temperature, and oxygen fugacity. *Contrib. Miner. Petrol.* **116**, 216–224 (1994).
59. Rousseeuw, P. J. & Van Zomeren, B. C. Unmasking multivariate outliers and leverage points. *J. Am. Stat. Assoc.* **85**, 633–639 (1990).
60. Rousseeuw, P. J. & Leroy, A. M. *Robust Regression and Outlier Detection* (Wiley, 2005).
61. Goodall, C. R. 13 Computation using the QR decomposition. (1993).
62. Gramatica, P. Principles of QSAR models validation: Internal and external. *QSAR Comb. Sci.* **26**, 694–701 (2007).
63. Amiri-Ramsheh, B., Zabihi, R. & Hemmati-Sarapardeh, A. Modeling wax deposition of crude oils using cascade forward and generalized regression neural networks: Application to crude oil production. *Geoenergy Sci. Eng.* **224**, 211613 (2023).

Author contributions

B.A.-R.: investigation, visualization, writing-original draft, M.N.A.: methodology, visualization, data analysis, M.S.: supervision, writing-review and editing. A.H.-S.: supervision, writing-review and editing.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.S. or A.H.-S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023