19th International Conference on Environmental Science and Technology

Kos, Greece, 3 to 6 September 2025

Advancing Temperature Swing Solvent Extraction Desalination with Machine Learning

CAIRONE S.[1,2,*], ZARRA T.[1], BELGIORNO V.[1], YIP N.Y.[2,3], NADDEO V.[1]

[1]Sanitary Environmental Engineering Division (SEED), Department of Civil Engineering, University of Salerno, Via Giovanni Paolo II #132, Fisciano, SA, 84084, Italy

[2]Department of Earth and Environmental Engineering, Columbia University, New York, New York 10027-6623, United States

[3]Columbia Water Center, Columbia University, New York, New York 10027-6623, United States

*corresponding author:
e-mail: scairone@unisa.it

Abstract The management of hypersaline wastewaters, including desalination brines, poses significant environmental and technical challenges. Temperature swing solvent extraction (TSSE) offers a promising solution by exploiting the temperature-dependent water affinity of low-polarity solvents, typically amines. In this study, machine learning (ML) is applied to enhance TSSE by predicting the mutual solubilities of water and amines. Experimental data and amine properties were used to train eight ML models. A stacking ensemble strategy yielded high predictive accuracy, with R^2 values of up to 97.8% for water-in-amine solubility and 95.8% for amine-in-water solubility. By integrating ML predictions with multi-objective optimization, the study identified ideal amines and operational temperatures.

Keywords: Artificial intelligence; Data-driven modelling; Process optimization; Solvent-driven desalination.

1. Introduction

Desalination is expanding rapidly worldwide, with global contracted capacity exceeding 109.2 million cubic meters per day (Borgomeo, 2024). However, desalination processes produce large volumes of highly concentrated saline solution (brine), comparable to the freshwater recovered. Managing desalination brine, as well as other hypersaline wastewaters like produced water from oil and gas operations, landfill leachate, and some industrial wastewaters, is challenging due to the technical limitations and potential negative environmental impacts of conventional disposing methods (Shah et al., 2022). Rich in both water and valuable minerals, these streams hold potential for sustainable resource recovery, enabling the production of fit-for-purpose desalinated water and the extraction of critical minerals. In this context, temperature swing solvent extraction (TSSE) has emerged as a promising technology (Boo et al., 2019; 2020; Shah et al., 2023), using low-polarity solvents with temperature-dependent water affinity (typically amines) to separate water from hypersaline solutions. Process efficiency depends on feed salinity, amine properties, and operating temperatures, with amine-water mutual solubilities playing a crucial role. This study applied machine learning (ML) to enhance TSSE by developing models that support the selection of optimal amines and operating temperatures.

2. Methodology

The IUPAC-NIST solubility data for aliphatic amines with water (Góral et al., 2012a; 2012b) served as the primary data source for this study. Data preprocessing involved deduplication, removal of irrelevant data, and outlier removal. Physicochemical properties of the amines were retrieved from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/) and incorporated into the dataset. Feature selection was performed using multicollinearity analysis to identify the most informative variables. By eliminating highly correlated variables, nine features were selected: operating temperature, molecular weight, partition coefficient, topological polar surface area, rotatable bond count, hydrogen bond donor count, hydrogen bond acceptor count, undefined atom stereocenter count, and molecular complexity. The final datasets consisted of 388 and 404 data points for the solubility of water in amines and amines in water, respectively. These datasets were randomly split, with 75% of data allocated for training and 25% for test. Adaptive boosting (AdaBoost), categorial boosting (CatBoost), extremely randomized trees (ExtraTrees), k-nearest neighbors (kNN), multilayer perceptron artificial neural network (MLP-ANN), random forest (RF), support vector regression (SVR), and extreme gradient boosting (XGBoost) were tested, with hyperparameters optimized via grid search. Model performance was assessed using mean squared error (MSE) and the coefficient of determination (R^2). All modeling tasks were conducted using Python. The best-peforming algorithms were combined using a stacking ensamble method, which combines predictions from multiple models to enhance overall accuracy. A multi-objective optimization was

carried out to identify optimal amines and operating temperatures. A trade-off strategy was employed to determine the most favorable solution that simultaneously: (i) maximizes the solubility of water in amine at low temperature (T_L), (ii) minimizes the solubility of water in amine at high temperature (T_H), and (iii) minimizes the solubility of amine in water at both T_L and T_H. Candidate T_L values ranged from 5 to 30 °C and T_H from 40 to 80 °C. The optimal low temperature (T_L*) and high temperature (T_H*) were determined individually per amine. The trade-off scores for water in amine (S_i) and amine in water (s_i) were calculated using Eqs. (1) and (2), respectively.

$$S_i = \max[X_i(T_L)] - \min[x_i(T_H)] \quad (1)$$

where $X_i(T_L)$ and $x_i(T_H)$ represent the solubility of water in the i-th amine at the T_L and T_H temperatures, respectively.

$$s_i = Y_i(T_L^*) + y_i(T_H^*) \quad (2)$$

where $Y_i(T_L^*)$ and $y_i(T_H^*)$ are the solubility of the i-th amine in water at T_L* and T_H* temperatures, respectively.

3. Results

All models demonstrated strong predictive performance for the solubility of water in amines, achieving an R^2 of up to 97.1% on the test set. The stacking approach further enhanced prediction accuracy, yielding the highest performance with an R^2 of 97.8% on the test set when combining SVR and CatBoost models. For predicting the solubility of amines in water, ensemble tree-based algorithms outperformed other models, reaching an R^2 of up to 94.3% on the test set. By integrating ExtraTrees, CatBoost, and XGBoost, the stacking ensamble improved performance further, achieving an R^2 of 95.8% on the test set. Based on the results of the proposed multi-objective optimization (**Figure 1**), N,N-Dimethylcyclohexylamine and N-Ethylcyclohexylamine were identified as the most suitable amines for the TSSE process.

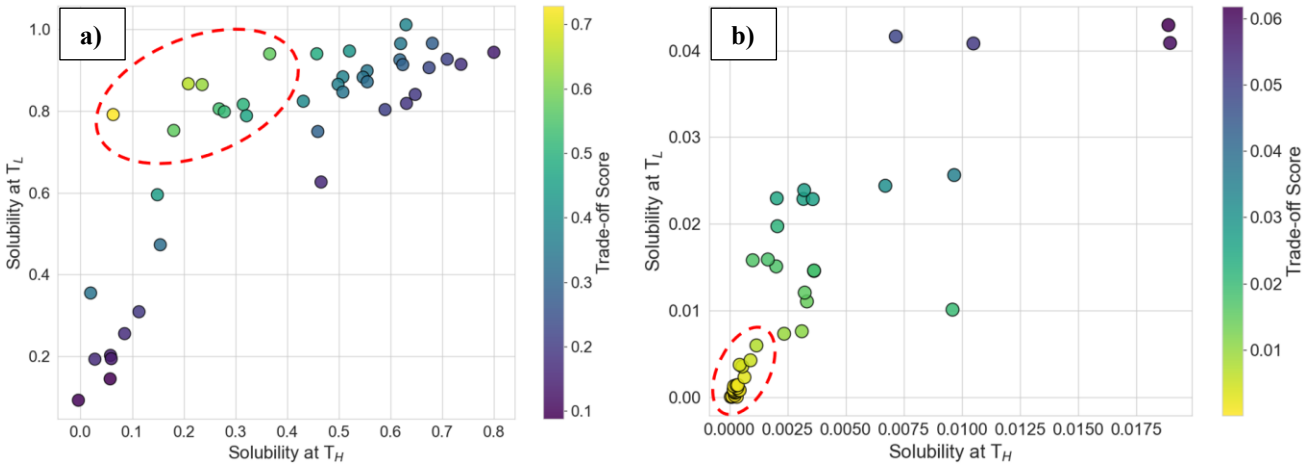

Figure 1. Trade-off scores for (a) water in amine (S_i) and (b) amine in water (s_i). The preferable trade-off scores are enclosed in the red closed curve.

4. Conclusion

ML models can accurately predict the mutual solubilities of water and amines. Integrating ML model outputs with multi-objective optimization provides a valuable tool for selecting optimal amines and operating temperatures for TSSE. Continued advancements in ML-enhanced TSSE have the potential to improve treatment efficiency and further expand the prospects for energy-efficient, cost-effective desalination of hypersaline brines.

References

Boo C., Winton R.K., Conway K.M. and Yip N.Y. (2019), Membrane-less and Non-Evaporative Desalination of Hypersaline Brines by Temperature Swing Solvent Extraction, *Environ. Sci. Technol. Lett.*, **6**, 359-364.

Boo C., Billinge I.H., Chen X., Shah K.M. and Yip N.Y. (2020), Zero Liquid Discharge of Ultrahigh-Salinity Brines with Temperature Swing Solvent Extraction, *Environ. Sci. Technol.*, **54**, 9124-9131.

Borgomeo E. (2024), A desalination research agenda fit for the 21st century, *Desalination*, **586**, 117910.

Góral M., Shaw D.G., Mączyński A., Wiśniewska-Gocłowska B. and Oracz P. (2012a), IUPAC-NIST Solubility Data Series. 96. Amines with Water Part 1. C_4–C_6 Aliphatic Amines, *J. Phys. Chem. Ref. Data*, **41**, 043106.

Góral M., Shaw D.G., Mączyński A., Wiśniewska-Gocłowska B. and Oracz P. (2012b), IUPAC-NIST Solubility Data Series. 96. Amines with Water Part 2. C_7–C_{24} Aliphatic Amines, *J. Phys. Chem. Ref. Data*, **41**, 043107.

Shah K.M., Billinge I.H., Chen X., Fan H., Huang Y., Winton R.K. and Yip N.Y. (2022), Drivers, challenges, and emerging technologies for desalination of high-salinity brines: A critical review, *Desalination*, **538**, 115827.

Shah K.M., Dach E., Winton R., Fan H. and Yip N.Y. (2023), Phase equilibria insights into amine-water-NaCl interactions in liquid-liquid biphasic systems for temperature swing solvent extraction desalination, *Desalination*, **548**, 116259.