TU Delft

Delft University of Technology

Machine Learning Approach for the Prediction of Eutectic Temperatures for Metal-Free Deep Eutectic Solvents

Lavrinenko, Anastasia K.; Chernyshov, Ivan Yu; Pidko, Evgeny A.

DOI
10.1021/acssuschemeng.3c05207

Publication date
2023

Document Version
Final published version

Published in
ACS Sustainable Chemistry and Engineering

Citation (APA)
Lavrinenko, A. K., Chernyshov, I. Y., & Pidko, E. A. (2023). Machine Learning Approach for the Prediction of Eutectic Temperatures for Metal-Free Deep Eutectic Solvents. *ACS Sustainable Chemistry and Engineering*, *11*(42), 15492-15502. https://doi.org/10.1021/acssuschemeng.3c05207

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

ACS Sustainable Chemistry & Engineering

pubs.acs.org/journal/ascecg

Open Access

This article is licensed under CC-BY 4.0

Research Article

Machine Learning Approach for the Prediction of Eutectic Temperatures for Metal-Free Deep Eutectic Solvents

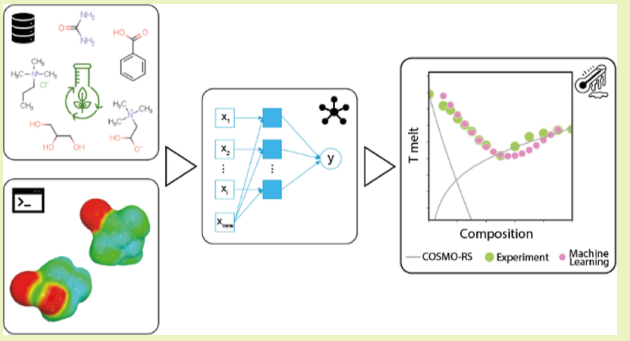
Anastasia K. Lavrinenko, Ivan Yu. Chernyshov, and Evgeny A. Pidko*

Cite This: ACS Sustainable Chem. Eng. 2023, 11, 15492−15502

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Deep eutectic solvents (DESs) represent an environmentally friendly alternative to conventional organic solvents. Their liquid range determines the areas of application, and therefore, the prediction of solid−liquid equilibrium (SLE) diagrams is essential for developing new DESs. Such predictions are not yet possible by using the current state-of-the-art computational models. Herein, we present an alternative model based on support vector regression integrating experimental data, a conductor-like screening model for real solvents simulations, and cheminformatic descriptors for predicting melting temperatures of binary metal-free DESs or ionic liquids, allowing the researcher to estimate the eutectic formation and SLE for specific combinations of components. The model was developed based on the manually collected database of 1648 mixture melting temperatures for 237 experimentally described DESs, and its accuracy was demonstrated by 5-fold cross-validation ($R^2 \sim 0.8$). The presented machine learning methodology empowers researchers to predefine the liquid range of the mixture and holds promise for efficient molecular combination screening, facilitating the discovery of tailored DESs for desired applications from catalysis and extraction to energy storage. By enabling a deeper understanding of DES behavior and the targeted design of these solvents, the proposed approach contributes to advancing green chemistry practices and to promoting sustainable solvent usage.



KEYWORDS: *deep eutectic solvents, COSMO-RS, machine learning, solid−liquid equilibrium, melting point prediction*

Downloaded via TU DELFT on November 20, 2023 at 12:00:14 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

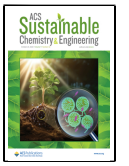
INTRODUCTION

Deep eutectic solvents (DESs) attracted considerable attention from the academic and industrial communities in the past decade due to their potential as often biomass-derived and biodegradable solvents with versatile and tunable characteristics. These materials are defined as a mixture of two or more components that, at a particular composition, present a high melting point depression in contrast to the initial components.[1,2] The depression in the melting point is associated with the formation of multiple hydrogen bonds between the DES components.[1]

Based on the structure, DESs were defined as a eutectic mixture of Lewis or Brønsted acids and bases with a general formula $Cat^+ X^- \cdot zY$ and classified into five main types presented in Figure 1.[3,4] Type III is the most experimentally and computationally studied DES[3,5] that consists of the quaternary ammonium or phosphonium salt and hydrogen bond donor (Figure 1). Types I, II, and IV DES involve metal salts and often contain non-nature-derived components, making them less attractive for future sustainable chemistry applications. DESs are generally described as a greener alternative to ionic liquids (ILs) due to their cheaper and easier synthesis, lower toxicity, high biodegradability, low melting point, nonflammability, thermal stability, and high dissolving ability.[6−10] However, we should acknowledge that the properties of DESs, including their toxicity and cost, strongly rely on their specific composition.[11] Some reports highlight the potential environmental toxicity and nongreen nature of DESs.[12−14] Nevertheless, the tunable characteristics of ILs and DESs established through their modular nature provide an opportunity for targeting sustainability as a design criterion. There is a growing commitment in the field to developing and synthesizing sustainable ILs and DESs, with a primary focus on reasonable selection of precursors.[14,15]

In this study, we consider both type III and type V DESs and conventional metal-free ILs regardless of their potential toxicity. This approach broadens the data set used for model creation, enhancing its robustness and applicability. While concerns about toxicity and biodegradability remain, especially when considering large-scale utilization, DESs undeniably contribute to the advancement of sustainable chemical

Received: August 15, 2023
Revised: September 26, 2023
Published: October 9, 2023



ACS Publications

© 2023 The Authors. Published by American Chemical Society

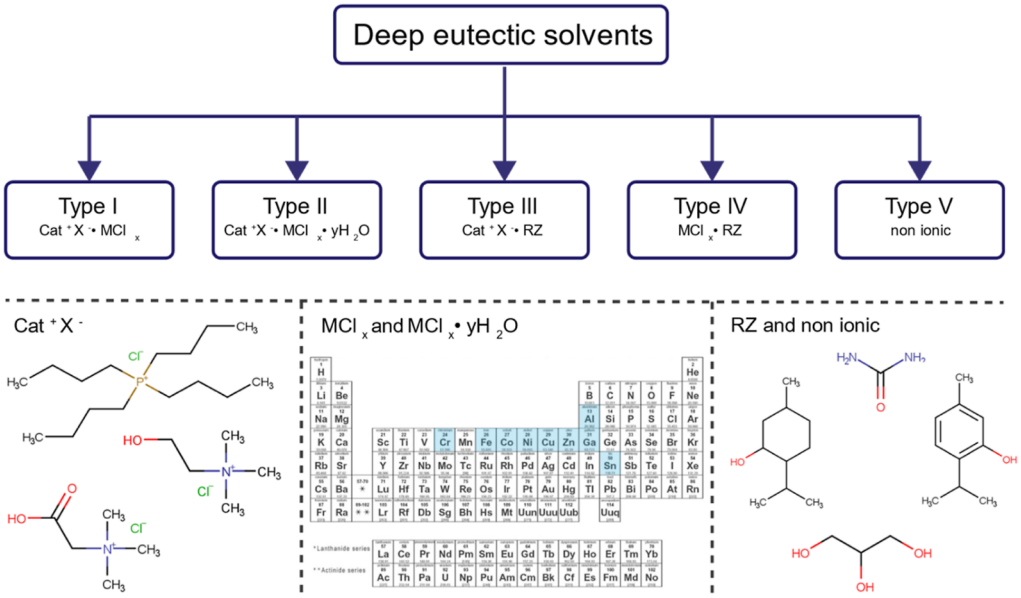


Figure 1. Classification of DESs based on the structures of the initial compounds.

processing. These solvents have already demonstrated their utility in enhancing the sustainability of such processes as chemical synthesis,[16−18] drug delivery,[19,20] biochemistry,[21−23] production of nanomaterials,[24,25] separation technologies,[26−30] electrochemistry,[31,32] and catalysis.[33]

So far, the development of new DES has relied primarily on tedious experimental search guided to a large extent by serendipity and the trial-and-error.[34,35] The efficiency of such an approach is further limited by the infinite number of combinations of the DES individual compounds and the practical challenge associated with the need to check wide ranges of pressures and temperatures in an iterative manner. Therefore, an alternative computational approach based on a universal theoretical model capable of calculating and predicting the characteristics of potential DES molecular combinations in silico is highly desirable. Such a model should be able to correctly predict the critical characteristics of DES, such as the liquid limits and the melting temperature.

Accurate prediction of the molecular properties in solution, as well as phase equilibria, is a challenging task due to the complex nature of the solvation process.[36] Among numerous approaches developed for the calculation of solvation energies,[37] the conductor-like screening model for real solvents (COSMO-RS) developed by Klamt et al.[38,39] stands out as the only model that is both able to account for the dependence of the solvation energy on the mixture composition and being fast enough for high-throughput computational screening. This established methodology is therefore chosen in this study as a representative state-of-the-art benchmark. Previous studies report the successful application of COSMO-RS to calculate phase diagrams and melting temperature in the case of nonideal systems and ILs.[40] However, COSMO-RS applicability for DES design is questionable as both successful[35,41] and hopeless cases were reported.[35,42] The application of group and group-interaction contribution methods was also proposed for melting temperature estimation and showed good accuracy on binary and ternary DESs.[43] These approaches assume that a property value is a linear function of the contribution of all functional groups in the compound's molecular structure; each functional group has the same contribution in every compound in which it appears. In addition, machine learning (ML) approaches have been successfully applied to describe the complex nature of the solvation process[44] and predict the properties of the molecules in the solid state.[45] The quantitative structure−property relationship (QSPR) model was reported to have good results in predicting the melting and freezing points of several DESs containing various hydrogen bond donors and chlorides.[46]

Herein, we propose a generic computational framework for rapidly screening molecular systems for DES design that can be established using a predictive model based on ML techniques. We analyzed the limitations and opportunities of the DES design using the COSMO-RS model. After the main reasons for the limited predictive power of such models were identified, the ML-based model for the estimation of the DES melting temperatures was constructed. Finally, the indirect approach to DESs' solid−liquid equilibria (SLE) prediction was established based on the trained ML model. Nowadays, when science moves toward more eco-friendly solutions, the methodology presented in this study becomes a valuable tool for researchers aiming to discover and develop novel DESs with reduced environmental impact. This accelerates the adoption of green chemistry practices and paves the way for a more sustainable future by providing efficient alternatives to traditional organic solvents.

MATERIALS AND METHODS

A unique database of 237 DESs with experimentally known eutectic temperature and composition was manually collected from the available literature. DES of type III, type V, and mixtures of IL in a total of 124 individual compounds were considered (Figure S1). The data set contains a simplified molecular-input line-entry system (SMILES), melting temperatures, fusion enthalpies of mixture compounds, experimental melting temperatures, and molar ratios of DESs (1648 points in total) (Table S1). In some cases, experimental mixture melting temperatures were not confirmed as eutectic temperatures. For such DESs, the lowest melting temperature was considered eutectic for future calculations. Fusion enthalpies for some DES compounds were not available from the literature (fusion enthalpies are known for both compounds for 208 DESs; fusion enthalpy is unknown for at least one compound for 29 DESs). DES

https://doi.org/10.1021/acssuschemeng.3c05207
ACS Sustainable Chem. Eng. 2023, 11, 15492−15502

compounds were classified as "component 1"—a compound with a lower individual melting temperature and "component 2"—a compound with a higher individual melting temperature.

Conductor-like screening model (COSMO) files for DES compounds were prepared for the predictive modeling of SLE diagrams by COSMO-RS. Initial geometries were generated from SMILES with the RDKit package (Version 2021.03.3),[47] subsequently optimized in an aqueous solution using the self-consistent reaction field method and used to calculate COSMO files by the Gaussian 16 software[48] at the BP86/TZVP level of theory. Two modeling approaches were used to generate the structures of the ionic compounds. COSMO files of the ionic pair were generated in the "CA" model, and COSMO files of separate cations and anions were generated in the "C + A" model (Figure S2).

The software package COSMOthermX (Version 19.0.5)[49] was used for the COSMO-RS SLE calculation. COSMO files and individual DES compound fusion properties (melting temperatures, fusion enthalpies) were incorporated into COSMOthermX; the BP_TZVP_C30_1701 parameterization was used. The SLE diagrams were calculated using the solid−liquid option in the 50−600 K temperature range and a temperature step of 8 K.

Several descriptors were used for ML modeling (Table S3). First, experimental descriptors (numbered 1−2) were collected from the literature, and chemoinformatic descriptors (numbered 3−26) were generated by the RDKit package from SMILES based on the two-dimensional (2D) or three-dimensional (3D) structures. Second, the descriptors based on COSMO-RS were computed by COSMOthermX with BP_TZVP_C30_1701 parameterization such as σ-moments (numbered 27−29), activity coefficients at infinite dilution (numbered 30), and integrated σ-profiles (numbered 31−37).

All ML calculations were conducted by the scikit-learn package (Version 0.24.2).[50] The input descriptors and the predicted target (melting temperatures) were normalized using a min−max scaling method

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

where x' is the scaled value, x is the initial value, and $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the column, respectively. The collected data were split randomly, and 80% of points were selected as the training set and 20% for the validation set. The "mixtures out" strategy was used in the current study.[51] Each unique DES mixture group was presented as a training or validation set. A similar tactic was used for cross-validation; each unique DES mixture group was included in the same fold. A training set with 5-fold cross-validation was used to select descriptors and the models' hyperparameter optimization, and a validation set was used to determine the accuracy of the final model. The random forest (RFR), gradient boosting (GBR), k-nearest neighbors (KNN), support vector machine (SVR), and multiple linear (MLR) regression algorithms, which are implemented in the scikit-learn package, were used for ML development. Optimization of hyperparameters was performed by using a grid search algorithm. Optimized parameters are presented in Table S5.

Three feature selection algorithms, namely, sequential forward selection (SFS), Gini indexes of random forest regression (GI), and maximal information coefficient (MIC), were used to represent the wrapper, embedded, and filter methods, respectively. The two-stage cascaded approach was performed for serial-based ensemble feature selection. The first stage feature selection was based on MIC; descriptors with coefficient value ≥0.2 were used as input data sets for the second selection stage based on GI, SFS, and MIC (more than 0.35). Since 5-fold cross-validation was used to select descriptors, each feature's frequency was calculated, and the average selection frequency was used as the threshold to obtain the final descriptors subset. Highly intercorrelated descriptors were determined by using the correlation matrix algorithm (Figure S3). Descriptors with an intercorrelation coefficient $R_{ij} > 0.7$ were recalculated as follows

$$\ln(\gamma) = \sum x_i \cdot \ln(\gamma_i^{\inf}) \tag{2}$$

$$\text{ChargeIndex} = \frac{\sum x_i \cdot ch_i^{\text{p}}}{\sum x_i \cdot a_i^{\text{p}}} \cdot \frac{\sum x_i \cdot a_i^{\text{n}}}{\sum x_i \cdot ch_i^{\text{n}}} \tag{3}$$

$$\text{SpherocityIndex} = \prod \text{SI}_i^{x_i} \tag{4}$$

$$\text{RadiusofGyration} = \sqrt{\frac{\sum \text{MW}_i \cdot \text{RG}_i^2}{\sum x_i \cdot \text{MW}_i}} \tag{5}$$

$$\text{PolarityIndex} = \frac{x_1 \cdot \int_{-0.4}^{0.4} \sigma\text{-profile}_1}{x_2 \cdot \int_{-0.4}^{0.4} \sigma\text{-profile}_2} \tag{6}$$

$$\text{SymmetricIndex\_HB} = \frac{\sum x_i \cdot \int_{-1.8}^{-1.1} \sigma\text{-profile}_i}{\sum x_i \cdot \int_{1.1}^{1.8} \sigma\text{-profile}_i} \tag{7}$$

$$\text{SymmetricIndex\_MF} = \frac{\sum x_i \cdot \int_{-1.1}^{-0.4} \sigma\text{-profile}_i}{\sum x_i \cdot \int_{0.4}^{1.1} \sigma\text{-profile}_i} \tag{8}$$

where x_i is a molar ratio, $\gamma_i^{\inf}$ is the activity coefficient at infinite dilution, ch_i^{p} and ch_i^{n} are the sum of positive and negative surface charges, respectively, a_i^{p} and a_i^{n} are the sum of positive and negative surface charges, respectively, SI_i is the spherocity index, MW_i is the molar weight, RG_i is the radius of gyration, and σ-profile$_i$ is σ-profile calculated by COSMO-RS of compound i. The final subset of selected descriptors is presented in Table S4.

The root mean squared error (RMSE), coefficient of determination (R^2), and mean absolute percentage error (MAPE) were used to reflect the performance of the prediction ability and are defined as follows

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (x_i - \hat{x}_i)^2} \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (x_i - \hat{x}_i)^2}{\sum_{i=1}^{n} (x_i - \bar{x}_i)^2} \tag{10}$$

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^{n} \frac{|x_i - \hat{x}_i|}{x_i} \tag{11}$$

where n is the number of data points, $\hat{x}_i$ is the predicted value of the ith sample, x_i is the corresponding actual value, $\bar{x}_i$ is the mean value of the experimental data set, and k is the number of descriptors used for prediction. The standard deviation associated with the mean of each metric over five random training/test splits is also provided.

RESULTS AND DISCUSSION

As fusion enthalpy is required to determine the eutectic point by the COSMO-RS model, a data set of DESs with known fusion enthalpies of both individual compounds (208 binary DESs) was used (Table S1). The COSMO-RS model calculated the SLE diagrams in the temperature range of 50−600 K with the BP_TZVP_C30_1701 parameterization as it was reported to have better results in eutectic point prediction among the three parameterizations BP_TZVP_C30_1601, BP_TZVP_C30_1701, and BP_TZVP_19.[35] The "CA" and "C + A" approaches were employed to model the ionic components. In the "CA" model, the ionic compound is represented as the ionic pair (Figure S2a), whereas in the "C + A" model, the ionic compound is represented by two isolated cation and anion species (Figure S2b).

https://doi.org/10.1021/acssuschemeng.3c05207
ACS Sustainable Chem. Eng. 2023, 11, 15492−15502

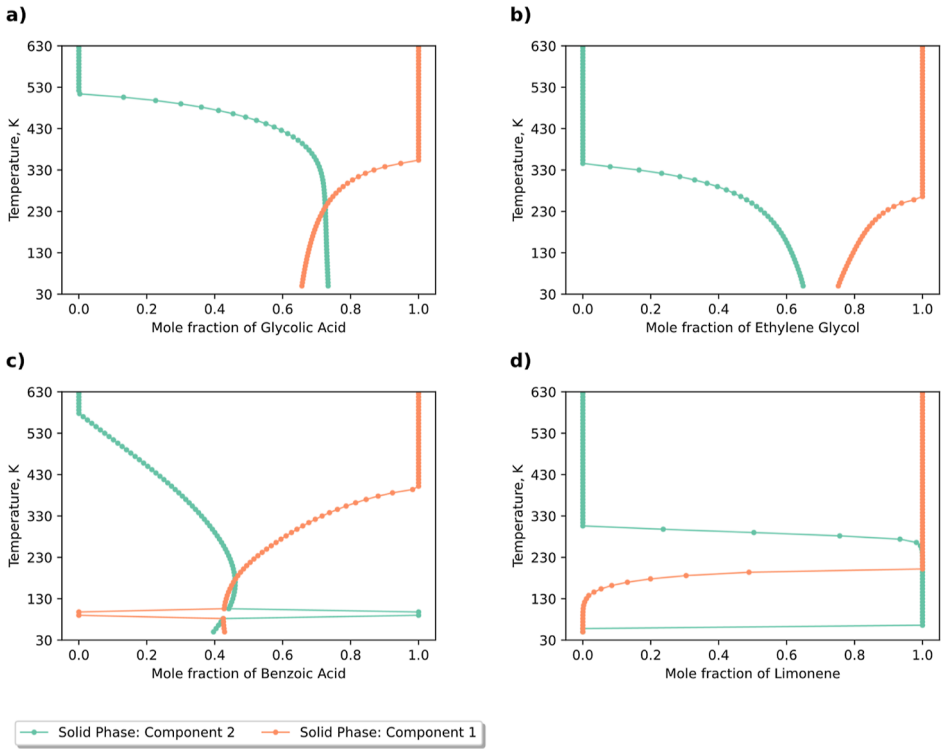


Figure 2. Representative SLE diagrams computed by the COSMO-RS approach with the "CA" model: (a) SLE of glycolic acid and betaine; (b) SLE of ethylene glycol and tetrabutylammonium chloride; (c) SLE of benzoic acid and choline chloride; and (d) SLE of limonene and capric acid.

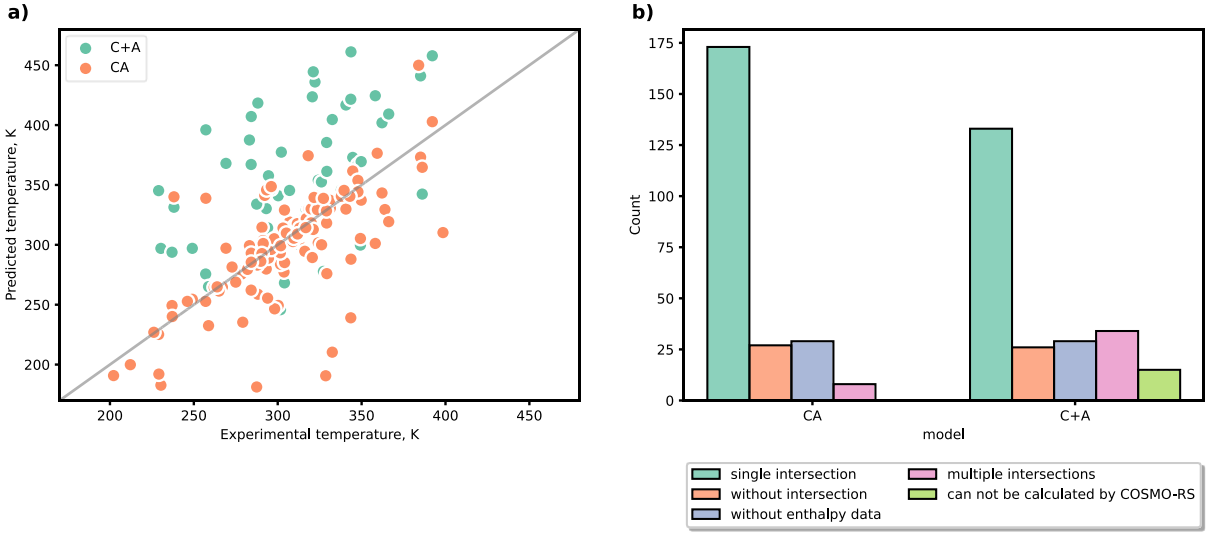


Figure 3. Comparison of the experimental and COSMO-RS-computed DES melting points (a) for "CA" and "C + A" model representations ($RMSE_{C+A}$ = 47.23, $RMSE_{CA}$ = 45.20) and (b) associated distribution of the number of intersections in the computed phase diagrams.

The COSMO-RS-computed eutectic points from phase diagrams of 208 binary DESs are summarized in Table S2. Three representative types of phase diagrams can be distinguished (Figure 2a–d). The diagrams featuring a single intersection allowed us to easily determine the eutectic melting temperatures (Figure 2a). In contrast, diagrams without the intersection (Figure 2b) and featuring multiple intersections (Figure 2c,d) could not be used to find the eutectic point. Even though for some diagrams with multiple intersections (Figure 2c), the eutectic point could be determined as the first intersection point, the overall reliability of the model in such cases is limited. Previous studies reported DES phase diagrams without intersection computed by COSMO-RS, which was attributed to activity coefficient underestimation.[35,41]

Comparison of the "CA" and "C + A" models revealed a broader applicability and a slightly better accuracy of the latter ionic pair approach in melting temperature prediction than the distinct ion approach (Figure 3). The "C + A" model also produced multiple intersections (such as illustrated with the SLE diagram in Figure 2d) more frequently than the "CA" model (Figure 3b). We also note that the "C + A" model failed in the calculation of IL mixtures where both compounds were represented by the "C + A" model (Figure 3b). The failures of the "C + A" model could be caused by the inaccurate

https://doi.org/10.1021/acssuschemeng.3c05207
ACS Sustainable Chem. Eng. 2023, 11, 15492−15502

description of the interactions between the highly charged ions within the current implementation of COSMO-RS.[52] Furthermore, the "CA" approach was previously reported to have better prediction ability than the "C + A" on the choline chloride-based eutectic mixtures.[41] Therefore, the "CA" representation has been selected for all further analyses and calculations.

The calculation of the SLE diagram with COSMO-RS is based on the experimental estimation of the fusion free energy.[52] The COSMO-RS model has several fundamental limitations including an incorrect description of the hydrogen bonding interactions of secondary and tertiary amines with solvents and electrostatic interactions of highly charged ions with unlocalized charge,[52] resulting in the observed low accuracy of the SLE prediction. Indeed, the best accuracy of eutectic point prediction by COSMO-RS was observed for DESs consisting of two nonionic compounds where the attraction between DES components is less pronounced ($RMSE_{CA}$ = 16.33). The difficulties in modeling IL by COSMO-RS related to the insufficient description of long-range interactions were reported previously.[53] The correction of the parameters corresponding to the hydrogen bond interaction in the model allows for a more accurate calculation of the eutectic temperatures.[42]

We propose to implement an ML-based approach to improve the accuracy of the DES melting temperature. We constructed an ML model based on a wide range of descriptors including melting temperatures of DES constituents and their molar ratio collected from literature, descriptors based on 2D and 3D molecular structures generated by the RDKit package, and descriptors based on the COSMO-RS model such as σ-moments, activity coefficients at infinite dilution, and integrated σ-profiles (Table S3). We added the mixture type to the descriptor subset as there is a statistically significant difference between the means of the melting temperatures (Figure S1c). We employed the ensemble feature selection approach to choose the best descriptor subset; combining multiple selection methods showed better results than using a single one.[54,55] The highly correlated descriptors were determined by Pearson's correlation coefficient (Figure S3) and descriptors with an intercorrelation coefficient $R_{ij} > 0.7$ were recalculated (eqs 2−8). The final subset of descriptors and their correlation with melting temperature are summarized in the Supporting Information (Table S4 and Figure S4). Correlation plots (Figure S4) showed that the selected descriptors allowed the differentiation between the type of DESs and melting temperature.

By definition, the melting point is the temperature where solid and liquid phases are in equilibrium, and it, therefore, can be expressed as the ratio of fusion enthalpy ΔH_{fus} and entropy change of melting ΔS_{fus}[56,57]

$$T_m = \frac{\Delta H_{fus}}{\Delta S_{fus}} \tag{12}$$

Fusion enthalpy is the energy required to transfer a substance from a solid to liquid. It depends on the crystal structure and intermolecular forces, while entropy represents an increase of disorder upon melting and depends on molecular symmetry and flexibility.[58] Indeed, interactions between molecules, their conformation, size, polarity, and surface charge distribution were the main factors influencing the melting point of organic molecules and ILs.[59−62]

The descriptors that we selected (Table S4) represent these factors. To account for the molecular interactions in the mixture, σ-profiles were calculated by the COSMO-RS. The feature selection algorithm suggested using integrals of σ-profiles 2 (from −1.8 to −1.1 e/nm^2), 3 (from −1.1 to −0.4 e/nm^2), 4 (from −0.4 to 0.4 e/nm^2), 5 (from 0.4 to 1.1 e/nm^2), and 6 (from 1.1 to 1.8 e/nm^2). Integral number 4 represents the polarity of compounds, integral numbers 3 and 5 represent electrostatic misfit intermolecular interaction in the mixture, while integrals 2 and 6 are related to hydrogen-bonding donor and acceptor capacities.[63] The symmetry of the σ-profile indicates favorable electrostatic and hydrogen-bonding interactions of the mixture with itself.[63] Therefore, the descriptors SymmetricIndex_MF and SymmetricIndex_HB (eqs 7 and 8) were introduced to reflect the symmetric shape of the mixture σ-profile. A strong dependence of the melting temperature of DES on both the introduced descriptors was observed (Figure S4). Notably, hydrogen bonding is hypothesized as the main intermolecular force in DES, affecting the melting point decrease.[64,65] To account for the nonideal nature of DESs,[64] we employed the activity coefficients at infinite dilution calculated by the COSMO-RS model representing the behavior of a single molecule surrounded by molecules of another compound.

The impact of molecular charge, size, and symmetry was also considered. The ChargeIndex descriptor represents the difference of positive and negative charges distributed over the surface to account for the electrostatic and size impact on the melting temperature. Values of charge and surface area correspond to polarizability and dipole moments, which affect intermolecular interactions and, thus, the melting point of the substance.[59] For example, for ILs, the dipole moment of the cation correlated to a decrease in fusion enthalpy and, thus, melting temperature.[66] The domination of the positive surface charge over the negative increased the melting temperature of DES (Figure S4). Such parameters like the SpherocityIndex, RadiusOfGyration, molecular weight, principal moment ratio, and InertialShapeFactor represent the influence of size, flexibility, and asymmetry of the DES components on the melting point. The correlation between the size and shape parameters of both the components and the mixture melting point of DESs was reported in the literature.[3,31,67] Interestingly, there is a strong impact of the geometrical parameters of component 2, which is mainly presented by the ionic components. The effect of the molecular shape of component 1 is less expressed. Similarly, the melting temperature of DES tended to decrease with the increasing molecular weight of component 2 (Figure S4). Indeed, larger sizes for ionic compounds led to weaker electrostatic interactions[67] and thus decreased lattice energy and melting temperature. In comparison, for nonionic molecules, larger size resulted in more expressed induced dipole interactions and higher melting temperatures.[59] Regarding the asymmetry of molecules, less symmetric and more flexible structures tend to have low lattice energy and melting temperature.[31,59] Indeed, DESs with a higher SpherocityIndex melted at higher temperatures (Figure S4).

The selected subset of descriptors (Table S4) was used to define the best ML model. For each model, we optimized hyperparameters via a grid search algorithm (Table S5) with 5-fold cross-validation and a "mixtures out" strategy.[51] The performance in melting temperature prediction for DES mixtures with various compositions of five ML models,

https://doi.org/10.1021/acssuschemeng.3c05207
ACS Sustainable Chem. Eng. 2023, 11, 15492−15502

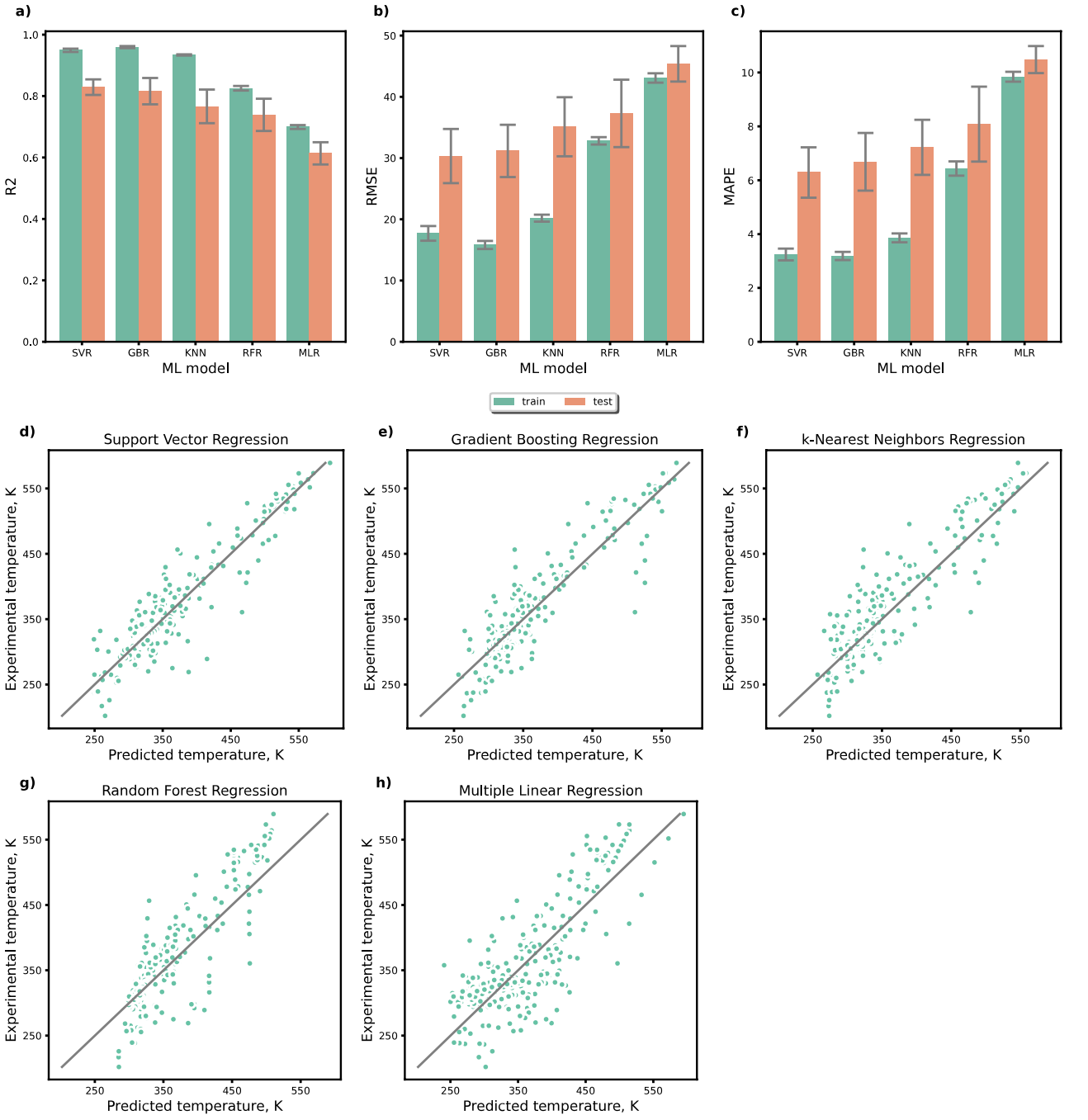


Figure 4. Comparison of performance of ML models in DES melting temperature prediction: (a) R^2; (b) RMSE; (c) MAPE; (d–h) comparison of experimental temperatures and predicted ones by SVR, GBR, KNN, RFR, and MLR, respectively.

namely, RFR, GBR, KNN, SVR, and MLR regressions, was also assessed using a 5-fold cross-validation approach combined with a "mixtures out" strategy. The cross-validated values of each metric with corresponding standard deviations are presented in Figure 4 and Table S6. SVR, GBR, and KNN performed best among the considered models (Table S6). Exceptionally high accuracy was observed for the SVR model, providing $R^2 \sim 0.85$, RMSE ~ 30−31 K, and MAPE ~ 5.57% using unseen (external) validation data. This prediction accuracy is comparable to the one obtained for the prediction of the melting point of organic molecules using natural language processing (RMSE = 36.88, $R^2 = 0.83$),[68] IL using the ML model (RMSE = 38.54, $R^2 = 0.76$),[69] DESs by QSPR (RMSE = 18.41, $R^2 = 0.80$),[46] and DESs using the group contribution method (MAPE = 5.67%).[43]

To accurately predict the eutectic point, the model should have good accuracy not only in mixture temperature estimation but also in predicting the correct correlation between the melting temperature and composition of the mixture. Herein, we attempted to estimate the eutectic temperature indirectly by using the developed models. We used the following algorithm based on the leave-one-out approach to investigate the performance of the developed models in eutectic point: (i) the ML model for mixture melting temperature prediction was trained on the data set with eliminated experimental points related to the unique DES

https://doi.org/10.1021/acssuschemeng.3c05207
ACS Sustainable Chem. Eng. 2023, 11, 15492−15502

mixture; (ii) for every unique DES mixture, melting temperature was predicted for the molar ratios in the range of 0.1−0.9 (16 points); (iii) the lowest temperature among the predicted ones and the corresponding molar ratio was considered as the predicted eutectic point. The predicted eutectic points are summarized in Table S2. Currently, SLE diagrams for DES are scarcely presented in the literature, and the formation of DES is not always confirmed by the experiments.[70−72] Some works reported investigating DESs at their supposed eutectic composition without justification by the phase diagram (Figure S1a),[64] which leads to an unknown real eutectic point and causes errors in the prediction models. Therefore, prediction accuracy was calculated based on the entries with the provided phase diagram in the literature.

All ML models showed much better performance in the eutectic temperature prediction than the COSMO-RS model (Table 1 and Figure 5), probably due to the ability of ML models to consider more complex nonlinear interactions of selected descriptors and melting temperature. Similarly, Koutsoukos et al. demonstrated that ML models predicted IL melting temperature better than the COSMO-RS model.[73] The best result was achieved using the SVR model (RMSE = 22.50, R^2 = 0.74) and showed a high potential in estimating the SLE diagram of DESs. However, all developed models had low R^2 values when predicting the eutectic composition (Table 1). It could be caused by the nonperfect strategy of considering the lowest temperature among the predicted ones and the corresponding molar ratio as a eutectic point, as it does not define the actual minima of the function. To better understand the spread of residuals and assess the reliability of the SVR model, we generated the predicted versus actual values plot and the residuals versus actual values plot (Figure S5). In the predicted versus actual values plot, we observed that most predicted values were closely scattered around the expected values, demonstrating a reasonably good fit of the SVR model supported by the corresponding evaluation metrics. The residual versus actual value plot indicated that the model did not exhibit any systematic bias in its predictions. The residuals appeared to be scattered randomly around zero, suggesting that the model's predictions were not consistently skewed in one direction. Most residuals fell within the RMSE value (dashed lines), supporting the conclusion that the model showed reasonable error limits. However, it is worth noting the presence of some outliers in each type of estimated DESs. These outliers were mainly observed when no phase diagrams were available from the literature. We retained such values in our analysis since the SVR model was trained on predicting the melting temperature of the mixture, independent of relying on eutectic point data. This highlights the model's ability to handle variations and deviations from established values, making it a valuable tool even when facing outliers not accounted for in the existing literature.

Table 1. Comparison of the Different Methodologies for the DES Eutectic Point Prediction

	eutectic temperature		eutectic composition	
model	R^2	RMSE	R^2	RMSE
COSMO-RS ("CA" model)	0.129	45.21	−0.561	0.18
GBR	0.683	24.95	0.041	0.14
RFR	0.434	33.31	−0.205	0.15
SVR	0.742	22.50	0.344	0.11
KNN	0.564	29.25	0.030	0.14

To compare the predictive power of the models for different DES types, we selected representative examples referring to the three types used in our research: IL mixture (Figure 5a), ionic/nonionic compounds (III-type) (Figure 5b), and nonionic/nonionic compounds (V-type) (Figure 5c,d). For all models, lower RMSE values were observed for DES of type III and V. The possible reason is that DES of type III is the most represented in our data set (Figure S1a). In the case of type V DESs, better accuracy could be caused by the less complex interactions and nonionic nature of the constituents. The result indicated that the types of DESs differ in the formation mechanism and that a large amount of SLE data related to other types of DES can help improve the prediction algorithm. The issues of transferability and accuracy in the model's applicability are well-recognized challenges in the field of ML for chemistry and materials science.[73,74] We observed that the developed models were less transferable in the case of entirely dissimilar molecules. In contrast, the transferability to the unseen DES was relatively high in the case of DESs with similar structures, which were commonly presented in the data set. For example, structures like sulfathiazole were less represented in our data set than IL, acids, alcohols, or phenols (Figure S1b); therefore, ML models had low accuracy in the prediction temperature of DES including sulfathiazole (Figure 5d). However, the SLE for DESs consisting of ammonium salts and/or carboxylic acids was estimated with high accuracy (Figure 5a−c). The developed ML model approach could be a much more efficient alternative to the COSMO-RS model for DES melting temperature estimation. Nonetheless, it is essential to continue research to address the challenges associated with model transferability and data set diversity, which will enhance the robustness and applicability of our approach.

Table 2 compares the results obtained from our prediction model (SVR) and those previously reported in the literature. One of the key differences in our study is the data set size we employed, which was at least twice as large as the data sets used in the reported methods. Our data set included III, V types of DESs and mixtures of IL, providing a broader and more diverse set of data points for analysis. Our prediction model demonstrated comparable results to the methods reported in the literature, even though a larger data set often introduces higher variability making achieving competitive performance challenging. Some methods achieved lower MAPE values. However, these methods were often tested on smaller data sets and were more focused on specific types of DESs. Interestingly, our model outperformed most other methods in predicting eutectic compositions. Such results suggest that our model has a particular strength in handling complex eutectic systems, highlighting its robustness and potential for broader applicability.

CONCLUSIONS

Summarizing, we have introduced an innovative approach that integrates publicly available molecular descriptors, SVR, and the physical principles of the COSMO-RS method. The developed framework is capable of predicting the melting temperatures of binary metal-free DESs or ILs, contributing to the rational design and application of these solvents.

The prediction model is based on the melting temperatures of initial compounds, their molar ratio, the quantity of

https://doi.org/10.1021/acssuschemeng.3c05207
ACS Sustainable Chem. Eng. 2023, 11, 15492−15502

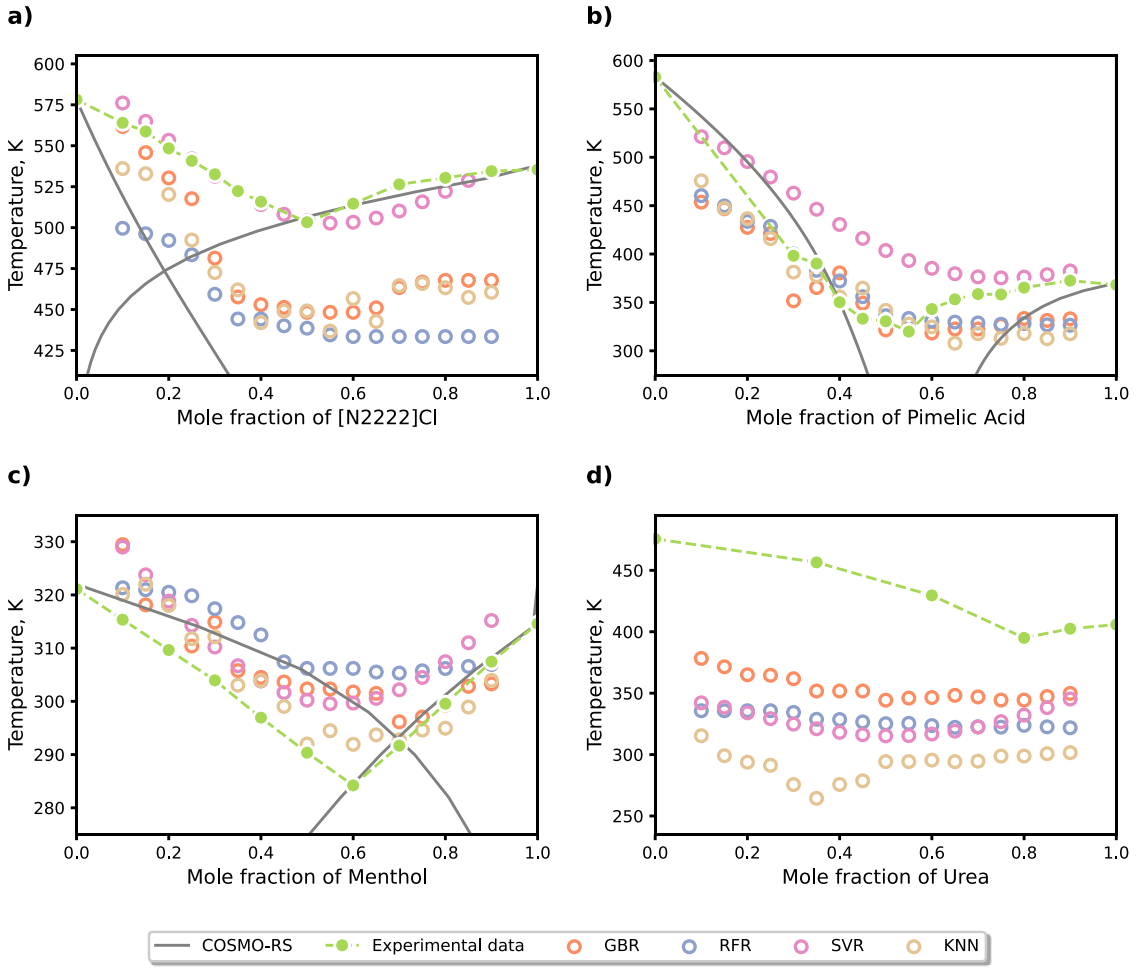


Figure 5. Prediction of DES's eutectic point by the ML approach. The comparison of the experimental and ML- and COSMO-RS-computed SLE diagrams for (a) tetraethylammonium chloride and choline chloride, (b) pimelic acid and tetrapropylammonium bromide, (c) menthol and phenylpropionic acid, and (d) urea and sulfathiazole; COSMO-RS calculations cannot be performed due to the unavailable fusion enthalpy of sulfathiazole.

Table 2. Comparison of Previously Published Methods for the DES Eutectic Point Prediction

method	type of DESs	number of estimated DESs	MAPE (eutectic temperature) (%)	MAPE (eutectic composition) (%)
SVR (our work)	type III, V, IL	237	7.27	17.79
COSMO-RS (TZVP-17)[35]	type III, V, IL	64	7.39	106.91
PC-SAFT[75]	L(−)-menthol/thymol + fatty acids	12	0.72	15.47
COSMO-RS (AB model)[41]	IL, choline chloride-based DES	23	8.87	
COSMO-SAC-10 (C + A method)[76]	type III, V, IL	162	6.63	27.18
regressed modified UNIFAC(a)[77]	type III, V	73	4.72	27.56
ElasticNet model and Redlich−Kister theory[78]	type III, V	15	3.75	38.84

deviation from an ideal mixture behavior, surface charge distribution, hydrogen bonding, and electrostatic and nonpolar intermolecular interactions, as well as geometrical parameters of size, flexibility, and asymmetry, which were identified as the key factors that are critical for the prediction of the DES mixture melting temperature. Several ML algorithms were considered. The best cross-validation accuracy in the melting temperature prediction of DESs was achieved for SVR (R^2 ~ 0.83, RMSE ~ 30 K), GBR (R^2 ~ 0.82, RMSE ~ 31 K), and KNN (R^2 ~ 0.77, RMSE ~ 35 K). The indirect approach of eutectic point estimation was proposed based on the developed models. We showed that the model based on the SVR could predict the SLE of DESs with the accuracy of eutectic temperature estimation of R^2 ~ 0.74, RMSE ~ 22−23 K, and eutectic composition estimation of R^2 ~ 0.34, RMSE ~ 0.11. Our approach bridges the gap between DES potential and practical application by offering a robust framework for predicting their behavior and guiding their rational design. The established ML approach enables the creation of environmentally friendly solvents tailored to specific needs while adhering to the principles of green chemistry and sustainability.

ML is currently the most straightforward and practical approach to predicting the eutectic point and estimating the

https://doi.org/10.1021/acssuschemeng.3c05207
ACS Sustainable Chem. Eng. 2023, 11, 15492−15502

SLE of DESs. However, to facilitate the development of such models, further efforts are necessary for the accurate systematization of the SLE experimental data, including the direct experimental search of eutectic point and the refinement of the data exclusively available in the graphical format.[79] Such systematization should significantly improve the predictive power of ML models and thus speed up the virtual screening of DESs.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssuschemeng.3c05207.

List of the used descriptors for ML model development and optimized hyperparameters (PDF)

Initial data set with SMILES strings, experimental melting points, and fusion enthalpies values (XLSX)

Results of eutectic temperature prediction by COSMO-RS and ML models (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Evgeny A. Pidko − *Inorganic Systems Engineering Group, Department of Chemical Engineering, Delft University of Technology, 2629 HZ Delft, The Netherlands;* orcid.org/0000-0001-9242-9901; Email: e.a.pidko@tudelft.nl

Authors

Anastasia K. Lavrinenko − *International Institute "Solution Chemistry of Advanced Materials and Technologies", ITMO University, Saint-Petersburg 191002, Russian Federation;* orcid.org/0000-0001-9863-8325

Ivan Yu. Chernyshov − *Inorganic Systems Engineering Group, Department of Chemical Engineering, Delft University of Technology, 2629 HZ Delft, The Netherlands*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssuschemeng.3c05207

Notes

The authors declare no competing financial interest.

The data supporting this study's findings are openly available at https://github.com/AstyLavrinenko/Eutectic-prediction.

The active research for this work has been completed in 2021.

■ ACKNOWLEDGMENTS

The use of the national computer facilities in this research was subsidized by NWO Domain Science. Anastasia Lavrinenko thanks support from ITMO University (Priority 2030 Federal Academic Leadership Program) and Nikita Serov for his valuable comments on the development of ML.

■ REFERENCES

(1) Zhang, H.; Lu, X.; González-Aguilera, L.; Ferrer, M. L.; del Monte, F.; Gutiérrez, M. C. Should Deep Eutectic Solvents Be Treated as a Mixture of Two Components or as a Pseudo-Component? *J. Chem. Phys.* **2021**, *154* (18), 184501.

(2) van den Bruinhorst, A.; Kollau, L. J. B. M.; Kroon, M. C.; Meuldijk, J.; Tuinier, R.; Esteves, A. C. C. A Centrifuge Method to Determine the Solid-Liquid Phase Behavior of Eutectic Mixtures. *J. Chem. Phys.* **2018**, *149* (22), 224505.

(3) El Achkar, T.; Greige-Gerges, H.; Fourmentin, S. Basics and Properties of Deep Eutectic Solvents: A Review. *Environ. Chem. Lett.* **2021**, *19* (4), 3397−3408.

(4) Singh, M. B.; Kumar, V. S.; Chaudhary, M.; Singh, P. A Mini Review on Synthesis, Properties and Applications of Deep Eutectic Solvents. *J. Indian Chem. Soc.* **2021**, *98* (11), 100210.

(5) Velez, C.; Acevedo, O. Simulation of Deep Eutectic Solvents: Progress to Promises. *WIREs Comput. Mol. Sci.* **2022**, *12* (4), No. e1598.

(6) Şahin, S. Tailor-Designed Deep Eutectic Liquids as a Sustainable Extraction Media: An Alternative to Ionic Liquids. *J. Pharm. Biomed. Anal.* **2019**, *174*, 324−329.

(7) Salehi, H. S.; Celebi, A. T.; Vlugt, T. J. H.; Moultos, O. A. Thermodynamic, Transport, and Structural Properties of Hydrophobic Deep Eutectic Solvents Composed of Tetraalkylammonium Chloride and Decanoic Acid. *J. Chem. Phys.* **2021**, *154* (14), 144502.

(8) Wang, J.; Zhang, S.; Ma, Z.; Yan, L. Deep Eutectic Solvents Eutectogels: Progress and Challenges. *Green Chem. Eng.* **2021**, *2* (4), 359−367.

(9) Liu, J.; Li, X.; Row, K. H. Development of Deep Eutectic Solvents for Sustainable Chemistry. *J. Mol. Liq.* **2022**, *362*, 119654.

(10) Xiong Chang, X.; Mujawar Mubarak, N.; Ali Mazari, S.; Sattar Jatoi, A.; Ahmad, A.; Khalid, M.; Walvekar, R.; Abdullah, E. C.; Karri, R. R.; Siddiqui, M. T. H.; Nizamuddin, S. A Review on the Properties and Applications of Chitosan, Cellulose and Deep Eutectic Solvent in Green Chemistry. *J. Ind. Eng. Chem.* **2021**, *104*, 362−380.

(11) Martins, M. A. R.; Pinho, S. P.; Coutinho, J. A. P. Insights into the Nature of Eutectic and Deep Eutectic Mixtures. *J. Solution Chem.* **2019**, *48* (7), 962−982.

(12) Jung, D.; Jung, J. B.; Kang, S.; Li, K.; Hwang, I.; Jeong, J. H.; Kim, H. S.; Lee, J. Toxico-Metabolomics Study of a Deep Eutectic Solvent Comprising Choline Chloride and Urea Suggests in Vivo Toxicity Involving Oxidative Stress and Ammonia Stress. *Green Chem.* **2021**, *23* (3), 1300−1311.

(13) Zaib, Q.; Eckelman, M. J.; Yang, Y.; Kyung, D. Are Deep Eutectic Solvents Really Green?: A Life-Cycle Perspective. *Green Chem.* **2022**, *24* (20), 7924−7930.

(14) Chen, Y.; Mu, T. Revisiting Greenness of Ionic Liquids and Deep Eutectic Solvents. *Green Chem. Eng.* **2021**, *2* (2), 174−186.

(15) Quintana, A. A.; Sztapka, A. M.; Santos Ebinuma, V. d. C.; Agatemor, C. Enabling Sustainable Chemistry with Ionic Liquids and Deep Eutectic Solvents: A Fad or the Future? *Angew. Chem.* **2022**, *134* (37), No. e202205609.

(16) Leron, R. B.; Wong, D. S. H.; Li, M.-H. Densities of a Deep Eutectic Solvent Based on Choline Chloride and Glycerol and Its Aqueous Mixtures at Elevated Pressures. *Fluid Phase Equilib.* **2012**, *335*, 32−38.

(17) Kamble, S. S.; Shankarling, G. S. Room Temperature Diazotization and Coupling Reaction Using a DES-Ethanol System: A Green Approach towards the Synthesis of Monoazo Pigments. *Chem. Commun.* **2019**, *55* (42), 5970−5973.

(18) Zhang, Q.; De Oliveira Vigier, K.; Royer, S.; Jérôme, F. Deep Eutectic Solvents: Syntheses, Properties and Applications. *Chem. Soc. Rev.* **2012**, *41* (21), 7108.

(19) Pedro, S. N.; Freire, M. G.; Freire, C. S. R.; Silvestre, A. J. D. Deep Eutectic Solvents Comprising Active Pharmaceutical Ingredients in the Development of Drug Delivery Systems. *Expet Opin. Drug Deliv.* **2019**, *16* (5), 497−506.

(20) Emami, S.; Shayanfar, A. Deep Eutectic Solvents for Pharmaceutical Formulation and Drug Delivery Applications. *Pharm. Dev. Technol.* **2020**, *25* (7), 779−796.

(21) Gorke, J. T.; Srienc, F.; Kazlauskas, R. J. Deep Eutectic Solvents for Candida Antarctica Lipase B-Catalyzed Reactions. *Ionic Liquid Applications: Pharmaceuticals, Therapeutics, and Biotechnology*; ACS, 2010; Vol. *1038*, pp 169−180.

(22) Kalhor, P.; Ghandi, K. Deep Eutectic Solvents for Pretreatment, Extraction, and Catalysis of Biomass and Food Waste. *Molecules* **2019**, *24* (22), 4012.

(23) Paiva, A.; Craveiro, R.; Aroso, I.; Martins, M.; Reis, R. L.; Duarte, A. R. C. Natural Deep Eutectic Solvents - Solvents for the 21st Century. *ACS Sustain. Chem. Eng.* **2014**, *2* (5), 1063−1071.

(24) Lee, J.-S. Deep Eutectic Solvents as Versatile Media for the Synthesis of Noble Metal Nanomaterials. *Nanotechnol. Rev.* **2017**, *6* (3), 271−278.
(25) Qu, Q.; Tang, W.; Tang, B.; Zhu, T. Highly Selective Purification of Ferulic Acid from Wheat Bran Using Deep Eutectic Solvents Modified Magnetic Nanoparticles. *Sep. Sci. Technol.* **2017**, *52* (6), 1022−1030.
(26) Shishov, A.; Bulatov, A.; Locatelli, M.; Carradori, S.; Andruch, V. Application of Deep Eutectic Solvents in Analytical Chemistry. A Review. *Microchem. J.* **2017**, *135*, 33−38.
(27) Rodriguez, N. R.; Gerlach, T.; Scheepers, D.; Kroon, M. C.; Smirnova, I. Experimental Determination of the LLE Data of Systems Consisting of {hexane + Benzene + Deep Eutectic Solvent} and Prediction Using the Conductor-like Screening Model for Real Solvents. *J. Chem. Thermodyn.* **2017**, *104*, 128−137.
(28) Bezold, F.; Weinberger, M. E.; Minceva, M. Assessing Solute Partitioning in Deep Eutectic Solvent-Based Biphasic Systems Using the Predictive Thermodynamic Model COSMO-RS. *Fluid Phase Equilib.* **2017**, *437*, 23−33.
(29) Liu, Y.; Friesen, J. B.; McAlpine, J. B.; Lankin, D. C.; Chen, S.-N.; Pauli, G. F. Natural Deep Eutectic Solvents: Properties, Applications, and Perspectives. *J. Nat. Prod.* **2018**, *81* (3), 679−690.
(30) Chemat; Vian, A.; Ravi; Khadhraoui; Hilali; Perino; Tixier. Review of Alternative Solvents for Green Extraction of Food and Natural Products: Panorama, Principles, Applications and Prospects. *Molecules* **2019**, *24* (16), 3007.
(31) Smith, E. L.; Abbott, A. P.; Ryder, K. S. Deep Eutectic Solvents (DESs) and Their Applications. *Chem. Rev.* **2014**, *114* (21), 11060−11082.
(32) Marcus, Y. *Deep Eutectic Solvents*; Springer International Publishing: Cham, 2019.
(33) Ünlü, A. E.; Arıkaya, A.; Takaç, S. Use of Deep Eutectic Solvents as Catalyst: A Mini-Review. *Green Process. Synth.* **2019**, *8* (1), 355−372.
(34) Bonab, P. J.; Esrafili, M. D.; Ebrahimzadeh, A. R.; Sardroodi, J. J. Molecular Dynamics Simulations of Choline Chloride and Phenyl Propionic Acid Deep Eutectic Solvents: Investigation of Structural and Dynamics Properties. *J. Mol. Graph. Model.* **2021**, *106*, 107908.
(35) Song, Z.; Wang, J.; Sundmacher, K. Evaluation of COSMO-RS for Solid-Liquid Equilibria Prediction of Binary Eutectic Solvent Systems. *Green Energy Environ.* **2021**, *6* (3), 371−379.
(36) Boobier, S.; Hose, D. R. J.; Blacker, A. J.; Nguyen, B. N. Machine Learning with Physicochemical Relationships: Solubility Prediction in Organic Solvents and Water. *Nat. Commun.* **2020**, *11* (1), 5753.
(37) Jalan, A.; Ashcraft, R. W.; West, R. H.; Green, W. H. Predicting Solvation Energies for Kinetic Modeling. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2010**, *106*, 211.
(38) Klamt, A. The COSMO and COSMO-RS Solvation Models. *WIREs Comput. Mol. Sci.* **2018**, *8* (1), No. e1338.
(39) Klamt, A.; Diedenhofen, M. A Refined Cavity Construction Algorithm for the Conductor-like Screening Model. *J. Comput. Chem.* **2018**, *39* (21), 1648−1655.
(40) Verma, N. R.; Gopal, G.; Anantharaj, R.; Banerjee, T. (Solid +liquid) Equilibria Predictions of Ionic Liquid Containing Systems Using COSMO-RS. *J. Chem. Thermodyn.* **2012**, *48*, 246−253.
(41) Abranches, D. O.; Larriba, M.; Silva, L. P.; Melle-Franco, M.; Palomar, J. F.; Pinho, S. P.; Coutinho, J. A. P. Using COSMO-RS to Design Choline Chloride Pharmaceutical Eutectic Solvents. *Fluid Phase Equilib.* **2019**, *497*, 71−78.
(42) Silva, L. P.; Fernandez, L.; Conceição, J. H. F.; Martins, M. A. R.; Sosa, A.; Ortega, J.; Pinho, S. P.; Coutinho, J. A. P. Design and Characterization of Sugar-Based Deep Eutectic Solvents Using Conductor-like Screening Model for Real Solvents. *ACS Sustain. Chem. Eng.* **2018**, *6* (8), 10724−10734.
(43) Hou, X.; Yu, L.; He, C.; Wu, K. Group and Group-interaction Contribution Method for Estimating the Melting Temperatures of Deep Eutectic Solvents. *AIChE J.* **2022**, *68* (2), No. e17408.
(44) Alibakhshi, A.; Hartke, B. Improved Prediction of Solvation Free Energies by Machine-Learning Polarizable Continuum Solvation Model. *Nat. Commun.* **2021**, *12* (1), 3584.
(45) Xin, D.; Gonnella, N. C.; He, X.; Horspool, K. Solvate Prediction for Pharmaceutical Organic Molecules with Machine Learning. *Cryst. Growth Des.* **2019**, *19* (3), 1903−1911.
(46) Khajeh, A.; Shakourian-Fard, M.; Parvaneh, K. Quantitative Structure-Property Relationship for Melting and Freezing Points of Deep Eutectic Solvents. *J. Mol. Liq.* **2021**, *321*, 114744.
(47) RDKit: Open-Source Cheminformatics. https://www.rdkit.org (accessed December, 2021).
(48) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16*, Revision C.01; Gaussian, Inc.: Wallingford CT, 2016.
(49) COSMOlogic GmbH & Co. KGa Dassault Systèmes Company. *COSMOtherm, Release 19*, 2019.
(50) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12* (85), 2825−2830.
(51) Muratov, E. N.; Varlamova, E. V.; Artemenko, A. G.; Polishchuk, P. G.; Kuz'min, V. E. Existing and Developing Approaches for QSAR Analysis of Mixtures. *Mol. Inf.* **2012**, *31* (3−4), 202−221.
(52) Klamt, A.; Eckert, F.; Arlt, W. COSMO-RS: An Alternative to Simulation for Calculating Thermodynamic Properties of Liquid Mixtures. *Annu. Rev. Chem. Biomol. Eng.* **2010**, *1* (1), 101−122.
(53) Brouwer, T.; Schuur, B. Model Performances Evaluated for Infinite Dilution Activity Coefficients Prediction at 298.15 K. *Ind. Eng. Chem. Res.* **2019**, *58* (20), 8903−8914.
(54) Bolón-Canedo, V.; Alonso-Betanzos, A. Ensembles for Feature Selection: A Review and Future Trends. *Inf. Fusion* **2019**, *52*, 1−12.
(55) Tsai, C.-F.; Sung, Y.-T. Ensemble Feature Selection in High Dimension, Low Sample Size Datasets: Parallel and Serial Combination Approaches. *Knowl. Base Syst.* **2020**, *203*, 106097.
(56) Alhadid, A.; Mokrushina, L.; Minceva, M. Design of Deep Eutectic Systems: A Simple Approach for Preselecting Eutectic Mixture Constituents. *Molecules* **2020**, *25* (5), 1077.
(57) Brown, R. J. C.; Brown, R. F. C. Melting Point and Molecular Symmetry. *J. Chem. Educ.* **2000**, *77* (6), 724.
(58) Godavarthy, S. S.; Robinson, R. L.; Gasem, K. A. M. An Improved Structure-Property Model for Predicting Melting-Point Temperatures. *Ind. Eng. Chem. Res.* **2006**, *45* (14), 5117−5126.
(59) Karthikeyan, M.; Glen, R. C.; Bender, A. General Melting Point Prediction Based on a Diverse Compound Data Set and Artificial Neural Networks. *J. Chem. Inf. Model.* **2005**, *45* (3), 581−590.
(60) Aguirre, C. L.; Cisternas, L. A.; Valderrama, J. O. Melting-Point Estimation of Ionic Liquids by a Group Contribution Method. *Int. J. Thermophys.* **2012**, *33* (1), 34−46.
(61) Bergström, C. A. S.; Norinder, U.; Luthman, K.; Artursson, P. Molecular Descriptors Influencing Melting Point and Their Role in Classification of Solid Drugs. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (4), 1177−1185.

https://doi.org/10.1021/acssuschemeng.3c05207
ACS Sustainable Chem. Eng. 2023, 11, 15492−15502

(62) Katritzky, A. R.; Jain, R.; Lomaka, A.; Petrukhin, R.; Maran, U.; Karelson, M. Perspective on the Relationship between Melting Points and Chemical Structure. *Cryst. Growth Des.* **2001**, *1* (4), 261−265.
(63) Eckert, F.; Klamt, A. Fast Solvent Screening via Quantum Chemistry: COSMO-RS Approach. *AIChE J.* **2002**, *48* (2), 369−385.
(64) Hansen, B. B.; Spittle, S.; Chen, B.; Poe, D.; Zhang, Y.; Klein, J. M.; Horton, A.; Adhikari, L.; Zelovich, T.; Doherty, B. W.; Gurkan, B.; Maginn, E. J.; Ragauskas, A.; Dadmun, M.; Zawodzinski, T. A.; Baker, G. A.; Tuckerman, M. E.; Savinell, R. F.; Sangoro, J. R. Deep Eutectic Solvents: A Review of Fundamentals and Applications. *Chem. Rev.* **2021**, *121* (3), 1232−1285.
(65) Sun, H.; Li, Y.; Wu, X.; Li, G. Theoretical Study on the Structures and Properties of Mixtures of Urea and Choline Chloride. *J. Mol. Model.* **2013**, *19* (6), 2433−2441.
(66) Rabideau, B. D.; Soltani, M.; Parker, R. A.; Siu, B.; Salter, E. A.; Wierzbicki, A.; West, K. N.; Davis, J. H. Tuning the Melting Point of Selected Ionic Liquids through Adjustment of the Cation's Dipole Moment. *Phys. Chem. Chem. Phys.* **2020**, *22* (21), 12301−12311.
(67) Xu, M.; Dou, H.; Peng, F.; Yang, N.; Xiao, X.; Tantai, X.; Sun, Y.; Jiang, B.; Zhang, L. Ultra-Stable Copper Decorated Deep Eutectic Solvent Based Supported Liquid Membranes for Olefin/Paraffin Separation: In-Depth Study of Carrier Stability. *J. Membr. Sci.* **2022**, *659*, 120775.
(68) Mi, W.; Chen, H.; Zhu, D.; Zhang, T.; Qian, F. Melting Point Prediction of Organic Molecules by Deciphering the Chemical Structure into a Natural Language. *Chem. Commun.* **2021**, *57* (21), 2633−2636.
(69) Low, K.; Kobayashi, R.; Izgorodina, E. I. The Effect of Descriptor Choice in Machine Learning Models for Ionic Liquid Melting Point Prediction. *J. Chem. Phys.* **2020**, *153* (10), 104101.
(70) Pereira, C. V.; Silva, J. M.; Rodrigues, L.; Reis, R. L.; Paiva, A.; Duarte, A. R. C.; Matias, A. Unveil the Anticancer Potential of Limomene Based Therapeutic Deep Eutectic Solvents. *Sci. Rep.* **2019**, *9* (1), 14926.
(71) Mjalli, F. S. Novel Amino Acids Based Ionic Liquids Analogues: Acidic and Basic Amino Acids. *J. Taiwan Inst. Chem. Eng.* **2016**, *61*, 64−74.
(72) Maugeri, Z.; Domínguez de María, P. Novel Choline-Chloride-Based Deep-Eutectic-Solvents with Renewable Hydrogen Bond Donors: Levulinic Acid and Sugar-Based Polyols. *RSC Adv.* **2012**, *2* (2), 421−425.
(73) Koutsoukos, S.; Philippi, F.; Malaret, F.; Welton, T. A Review on Machine Learning Algorithms for the Ionic Liquid Chemical Space. *Chem. Sci.* **2021**, *12* (20), 6820−6843.
(74) Liu, X.; Yu, M.; Jia, Q.; Yan, F.; Zhou, Y.-N.; Wang, Q. Leave-One-Ion-out Cross-Validation for Assisting in Developing Robust QSPR Models of Ionic Liquids. *J. Mol. Liq.* **2023**, *388*, 122711.
(75) Martins, M. A. R.; Crespo, E. A.; Pontes, P. V. A.; Silva, L. P.; Bülow, M.; Maximo, G. J.; Batista, E. A. C.; Held, C.; Pinho, S. P.; Coutinho, J. A. P. Tunable Hydrophobic Eutectic Solvents Based on Terpenes and Monocarboxylic Acids. *ACS Sustain. Chem. Eng.* **2018**, *6* (7), 8836−8846.
(76) Peng, D.; Alhadid, A.; Minceva, M. Assessment of COSMO-SAC Predictions for Solid-Liquid Equilibrium in Binary Eutectic Systems. *Ind. Eng. Chem. Res.* **2022**, *61* (35), 13256−13264.
(77) Song, Z.; Chen, J.; Qin, H.; Qi, Z.; Sundmacher, K. Extending UNIFAC Models for Solid-Liquid Equilibria Prediction and Design of Eutectic Solvent Systems. *Chem. Eng. Sci.* **2023**, *281*, 119097.
(78) Wang, R.; Chen, J.; Song, Z.; Qi, Z. Bridging Machine Learning and Redlich-Kister Theory for Solid-Liquid Equilibria Prediction of Binary Eutectic Solvent Systems. *Ind. Eng. Chem. Res.* **2023**, *62* (12), 5382−5393.
(79) Chirico, R. D.; de Loos, T. W.; Gmehling, J.; Goodwin, A. R. H.; Gupta, S.; Haynes, W. M.; Marsh, K. N.; Rives, V.; Olson, J. D.; Spencer, C.; Brennecke, J. F.; Trusler, J. P. M. Guidelines for Reporting of Phase Equilibrium Measurements (IUPAC Recommendations 2012). *Pure Appl. Chem.* **2012**, *84* (8), 1785−1813.

https://doi.org/10.1021/acssuschemeng.3c05207
ACS Sustainable Chem. Eng. 2023, 11, 15492−15502