Systems & Control Transactions

Research Article - Peer Reviewed Conference Proceeding
ESCAPE 35 - European Symposium on Computer Aided Process Engineering
Ghent, Belgium. 6-9 July 2025
Jan F.M. Van Impe, Grégoire Léonard, Satyajeet S. Bhonsale,
Monika E. Polańska, Filip Logist (Eds.)

PSE PRESS

Systematic comparison between Graph Neural Networks and UNIFAC-IL for solvent pre-selection in liquid-liquid extraction

Edgar Ivan Sanchez Medina[a]*, Ann-Joelle Minor[a], and Kai Sundmacher[a,b]*

[a] Max Planck Institute for Dynamics of Complex Technical Systems, Department of Process Systems Engineering, Magdeburg, Saxony-Anhalt, Germany
[b] Otto von Guericke University Magdeburg, Chair of Process systems Engineering, Magdeburg, Saxony-Anhalt, Germany
* Corresponding Authors: sanchez@mpi-magdeburg.mpg.de , sundmacher@mpi-magdeburg.mpg.de

ABSTRACT

Solvent selection is a critical decision-making process that balances economic, environmental, and societal factors. The vast chemical space makes evaluating all potential solvents impractical, necessitating pre-selection strategies to identify promising candidates. Predictive thermodynamic models, such as the UNIFAC model, are commonly used for this purpose. Recent advancements in deep learning have led to models like the Gibbs-Helmholtz Graph Neural Network (GH-GNN), which overall offers higher accuracy in predicting infinite dilution activity coefficients over a broader chemical space than UNIFAC. This study presents a systematic comparison of solvent pre-selection using GH-GNN and UNIFAC-IL in the context of liquid-liquid extraction. The original GH-GNN model is extended to simultaneously predict organic and ionic systems. This extended GH-GNN model predicts more than 92 % of the logarithmic IDACs with an absolute error of less than 0.3. By comparison, UNIFAC-based models only achieve such accuracy for less than 65 %. A case study is used involving the ionic liquid ethyl-3-methylimidazolium tetrafluoroborate ([EMIM][BF4]) and caprolactam, relevant for the solvolytic depolymerization of polyamide 6. Results indicate a significant correlation in solvent rankings across both methods, with a Spearman's coefficient of 0.62, suggesting that deep learning-based models like GH-GNN are viable alternatives for solvent pre-selection. Additionally, chemical similarity metrics, such as Tanimoto similarity, can assess confidence in solvent rankings, allowing users to determine acceptable risk levels in predictions across a vast chemical space.

Keywords: Artificial Intelligence, Machine Learning, Phase Equilibria, Polymers

INTRODUCTION

Solvents are used in various areas of chemical processes, including reaction and separation units. Selecting appropriate solvents is a complex task that involves decision-making from multiple perspectives. On one hand, the actual thermodynamic performance of the solvents must be considered. On the other hand, the implications of using such solvents must also be considered from environmental and societal perspectives.

From this perspective, the task of predicting solvent properties becomes central to assessing the overall performance of solvents within a specific process system. While important molecular properties can be measured in the laboratory, the vast extension of chemical space makes it inefficient to explore all promising solvent alternatives solely through experimental means. Therefore, predictive thermodynamic models are of key interest during the early stages of process design, especially for those involving solvents.

Pre-selection strategies aim to reduce the time and resources needed during the design of chemical processes by exploring the chemical space and allowing to focus on the most promising solvent candidates. These candidates can then be evaluated using more precise methods and eventually, the top-performer(s), could be investigated experimentally. Typically, the assessment of solvent performance in extraction applications relies on

the prediction of activity coefficients at infinite dilution (IDACs) [1]. Under these conditions, the solvent effect is at its maximum, which is why it is commonly used as a starting point for evaluating and ranking solvent candidates.

In this context, a popular choice is to use the UNIFAC model. However, with the rise of deep learning, novel alternative models have recently been proposed for predicting IDACs, showing overall more accurate predictions [2,3]. One of these alternatives is the so-called Gibbs-Helmholtz Graph Neural Network (GH-GNN), which was originally trained to predict IDACs of organic systems [4], and has also been extended to handle polymer solutions [5].

While the GH-GNN model has been previously used for solvent pre-selection [1], a systematic comparison of the resulting pre-selection to traditional methods is still lacking. Furthermore, given that the GH-GNN model was previously limited to predicting organic systems, we present here a series of extension strategies for expanding the GH-GNN model to simultaneously predict organic systems and mixtures containing ionic liquids (ILs).

This extension is particularly relevant in the context of chemical recycling of polyamide 6, also known as nylon 6, where ionic liquids have been identified as promising candidates due to their non-volatility and catalytic properties, though a complete assessment and process simulation were previously hindered by missing thermodynamic data [8]. The advantage of this approach is elucidated regarding the separation of the monomer caprolactam from the ionic liquid ethyl-3-methylimidazolium tetrafluoroborate ([EMIM][BF4]).

This work is structured as follows. First, the GH-GNN model is briefly described. Then, the extension strategies for systems containing ionic-liquids are presented. Subsequently, using the aforementioned case study, a systematic comparison of solvent pre-selection is conducted by comparing the extended GH-GNN model to UNIFAC-IL. Finally, conclusions and future directions are discussed.

METHODS

UNIFAC-IL

The UNIFAC model is a method used to estimate the excess Gibbs energy of a system, which in turn is used to determine the activity coefficients of its components. It operates on the principle of a "solution of groups," where molecules are decomposed into functional groups. This approach simplifies the model by reducing the parameter space from one that requires individual parameters for each unique molecule to one that only needs parameters for each type of group. This reduction in complexity enhances the model's predictive capabilities, allowing it to explore a wider range of chemical systems more efficiently.

Originally, the UNIFAC model was parameterized for predicting organic mixtures. However, extensions of the UNIFAC model to systems that include ionic liquids have also been developed in the past. In this work, we use the UNIFAC-IL extension from [6]. In this extension, a neural recommender system was first trained on experimentally determined IDACs of mixtures containing ionic liquids. Then, this recommender system was used to fill in the missing entries in a previously reported solute-solvent interaction matrix. Finally, the completed matrix was used to regress the binary interaction parameters of the UNIFAC model extended to ionic liquids.

This version of UNIFAC-IL [6] employs the original form of the UNIFAC model, and contains all necessary binary interaction parameters for 22 conventional main groups, 11 cationic main groups and 38 anionic main groups, with overall 102 subgroups. For further details on the model, readers are referred to the original publications mentioned earlier.

Gibbs-Helmholtz Graph Neural Network

The GH-GNN model was first introduced in [4]. Unlike the UNIFAC model, the GH-GNN model uses a graph representation of molecules, where non-hydrogen atoms are represented as nodes and covalent chemical bonds as edges. Each node and edge in the molecular graph is assigned a vector of binary values containing relevant attributes of the atom or bond, respectively. Information about hydrogen atoms is implicitly included as part of the node features in the graph.

The GH-GNN model consists of two graph neural networks (GNNs) operating at different levels. The first GNN processes molecular graphs to obtain learned representations tailored for the prediction of IDACs. The second GNN operates at the mixture level, using a graph where the learned molecular representations of distinct chemical species in the mixture define the nodes and hydrogen-bonding interactions are represented as edges. These two GNNs are sequentially arranged to predict parameters $K_{1,ij}$ and $K_{2,ij}$ of the following equation derived from the Gibbs-Helmholtz relationship:

$$\ln \gamma_{ij}^{\infty} = K_{1,ij} + \frac{K_{2,ij}}{T} \quad (1)$$

Parameters $K_{1,ij}$ and $K_{2,ij}$ are temperature-independent and specific to the precise solvent-solute combination. As shown in Eq. 1, predicting these two parameters is sufficient to estimate the temperature-dependent IDACs. For further details on the model, readers are referred to the original publication [4].

Tanimoto similarity for risk level assessment

As demonstrated in the original publication [4], the accuracy of the GH-GNN predictions is correlated with the distance within the chemical space between the

systems observed during training and the systems being predicted. If the distance is relatively short, the predictions tend to be more accurate. Conversely, as the distance increases, the accuracy of the predictions begins to decrease.

The distance between distinct solute-solvent systems can be approximated using the Tanimoto similarity metric, as mentioned in [4]. This metric is designed to estimate the "level of extrapolation" that the GH-GNN must undertake when predicting a system containing a chemical species that was not observed during training (extrapolated species).

The Tanimoto similarity metric is computed by taking the Jaccard distance between the molecular fingerprints of the "extrapolated species" and the set of the n most similar molecules observed during training. This allows the model user to estimate the level of accuracy of the GH-GNN predictions and can inform decisions regarding the level of risk the user is willing to take during the exploration of the chemical space. This feature is not available for models like UNIFAC.

Solvent pre-selection for liquid-liquid extraction

Measuring the thermodynamic performance of solvents participating in liquid-liquid extraction processes is typically carried out by computing the separation factor [9] defined by Eq. 2

$$K_s = \frac{\gamma_i^R \gamma_j^E}{\gamma_i^E \gamma_j^R} \quad (2)$$

In Eq. 2, the effects of the solvent in the raffinate phase R and the extract phase E are considered. In the case study analyzed here, species i corresponds to the monomer caprolactam, and species j corresponds to the ionic liquid [EMIM][BF4]. Therefore, the ratio γ_i^R/γ_i^E measures the performance of the solvent within the extraction step, where it is desirable for caprolactam in the extract phase to be as enriched as much as possible while its content in the raffinate phase to be as low as possible. Similarly, the ratio γ_j^E/γ_j^R measures the performance at the recovery step, where [EMIM][BF4] should be easily separable from the extract phase and preferably move to the raffinate phase.

As a pre-selection metric, the following separation factor at infinite dilution was used, which similarly to Eq. 2 considers the solvent performance at both the extraction and recovery stages:

$$K_s^\infty = \frac{\gamma_{sj}^\infty}{\gamma_{is}^\infty} \quad (3)$$

Eq. 3 results in larger separation factor values for solvents that are easily separable from the raffinate phase (mostly composed by [EMIM][BF4]), while entraining caprolactam as much as possible from the original mixture. This separation factor at infinite dilution was computed for every solvent considering the ambient temperature of 298.15 K, on which the extraction process is assumed to be operated.

RESULTS

GH-GNN extension strategies to ionic liquids

Three different strategies were studied to extend the original GH-GNN model from organic systems to ionic liquids. For this purpose, experimental IDAC data reported in the literature for ionic systems was included alongside the data for organic systems originally used to develop the GH-GNN [4]. The ionic liquids dataset corresponds to the compendium of ILThermo (v2.0), originally retrieved by [6] and further processed by [3].

The first extension strategy (S1) involved a complete re-training of the GH-GNN framework using only ionic liquid data. The second strategy (S2) consisted of a pre-training phase of the model using only organic systems, followed by a fine-tuning phase with the ionic liquid data. The third strategy (S3) involved a concurrent re-training of the GH-GNN model using the combined datasets for both organic and ionic liquid systems. In all cases, the same training and test split was used as in the original referenced publications [3,4].

Table 1 compares the three strategies based on the percentage of systems in the test set that are predicted with an absolute error of less than or equal to 0.3.

Table 1: Comparison of the percentage of systems predicted with an absolute error less than or equal to 0.3.

Model	Organic	Ionic liquid
(Original) GH-GNN	91.7 %	16.8 %
S1: (IL) GH-GNN	10.2 %	94.3 %
S2: (pre-trained) GH-GNN	10.9 %	94.1 %
S3: (concurrent) GH-GNN	92.4 %	94.0 %
GNN [3]		93.8 %
Matrix Completion [3]		94.1 %
UNIFAC (Dortmund)	64.3 %	
UNIFAC-IL [6]		50.5 %

It can be observed that while the original GH-GNN model excels in predicting IDACs for organic systems, its performance when extrapolating to systems with ionic liquids is poor, and is considerably worse than UNIFAC-IL. In contrast, the model developed using the first strategy S1 (i.e., (IL) GH-GNN) achieves high accuracy for ionic liquid systems, even surpassing the GNN model previously reported by [3]. However, when tested for predicting organic systems, the model performs poorly. Analyzing the GH-GNN model extended with the second strategy S2 (i.e., (pre-trained) GH-GNN), the accuracy for predicting organic systems slightly increases, but at the expense of slightly decreasing the accuracy for ionic

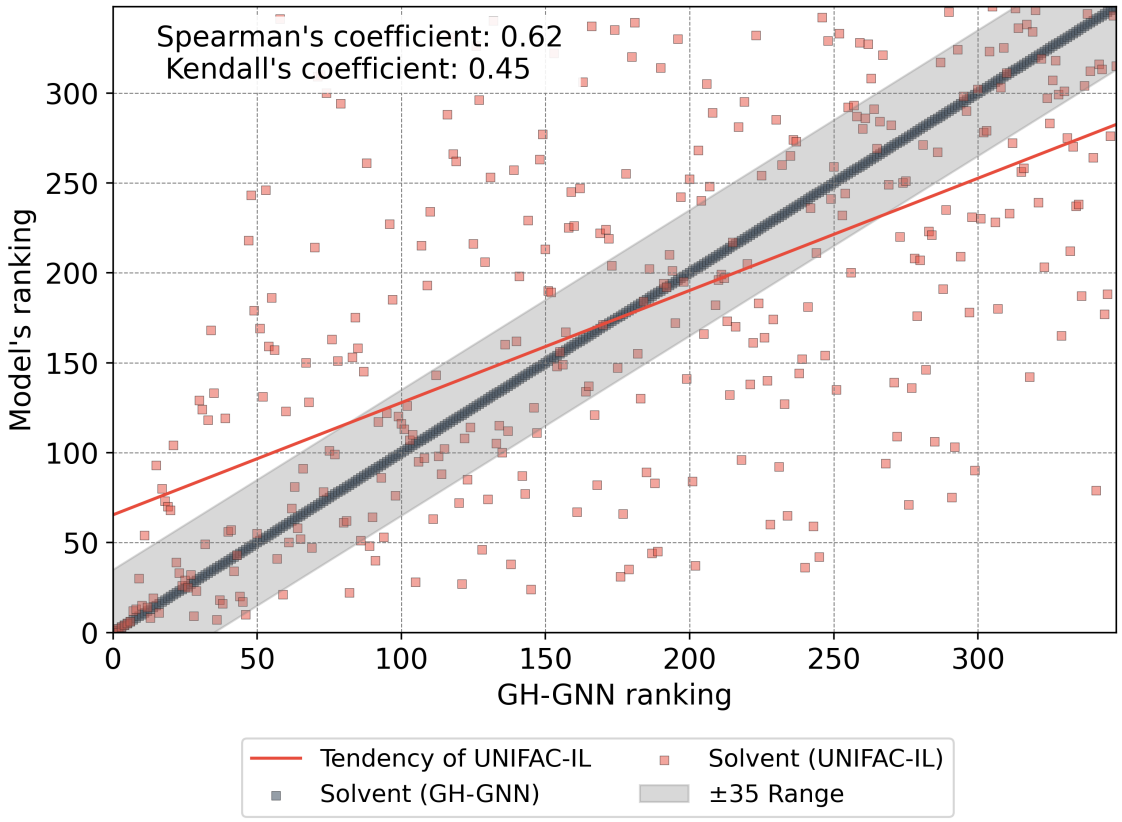

Figure 1: Comparison of relative solvent rankings obtained by the GH-GNN model and the UNIFAC-IL model. The linear tendency of the ranking of UNIFAC-IL with respect to the one of GH-GNN is also shown. The gray shaded area indicates ± 35 (10 %) places in the ranking of the GH-GNN.

liquid systems. This mutually exclusive behaviour in accurately predicting either organic or ionic systems highlights the difficulty of exploring vast chemical spaces while extrapolating beyond the types of systems used during training.

Remarkably, the GH-GNN model extended using the third strategy S3 (i.e., (concurrent) GH-GNN) achieves the highest accuracy for predicting organic systems among all compared models, while also performing remarkably well in predicting systems with ionic liquids. Interestingly, as a result of including ionic liquid data, the model's performance in predicting organic systems even surpasses that of the original GH-GNN model. This suggests that the model is able to extract relevant information from systems of organic molecules infinitely diluted in ionic liquids, which is beneficial for capturing effects in organic systems alone. The comparison of the extended (concurrent) GH-GNN model with the UNIFAC models demonstrates the potential of this alternative methods to predict IDACs with significantly higher overall accuracy.

Comparison of solvent pre-selection between GH-GNN and UNIFAC-IL

The GH-GNN model extended to ionic liquids using the concurrent strategy was employed here to compare solvent pre-selection against UNIFAC-IL. The selection of solvents is aimed at separating caprolactam from ([EMIM][BF4]), a mixture arising in the context of the chemical recycling of polyamide 6. In total, 700 solvents were considered as potential candidates, based on the selection of molecules collected by [7]. Only organic molecules are included among the candidate solvents.

The chemical space that UNIFAC-IL can predict is more restricted compared to that of the GH-GNN model. As a result, the first difference between the methods is encountered: while the necessary IDACs involving all 700 solvents can be predicted by GH-GNN, only the IDACs for 359 solvents are feasible to predict with UNIFAC-IL due to limitations in the "set of functional groups" representation of the molecules and/or the lack of required group binary interaction parameters.

Fig. 1 presents the comparison of the obtained solvent rankings according to GH-GNN and UNIFAC-IL. A Spearman's rank correlation coefficient of 0.62 is obtained between the two rankings, indicating a moderate positive monotonic correlation. When computing Kendall's correlation coefficient for the rankings, a value of 0.45 is obtained, also suggesting a moderate positive correlation. This implies that while the two methods do not rank the solvents identically, there is a substantial level of agreement in their rankings, indicating that they may be capturing similar underlying patterns or properties of the solvents.

However, the exact solvent rankings obtained using GH-GNN and UNIFAC-IL also show significant discrepancies that might influence which solvents are preferred over others. Given that the accuracy of GH-GNN for computing IDACs is significantly higher than that of UNIFAC-IL, it is reasonable to assume that the solvent ranking provided by GH-GNN offers a more reliable description of the physical behaviour.

In total, 42.4 % of the solvents lie within the 10 % (i.e., 35 ranking places) deviation from each method (shaded area in Figure 2). Interestingly, both methods agree on the set of top 7 solvents, which are shown in Table 2. Note that the ranking of these solvents is solely based on their performance according to Eq. 3. Other practical aspects relevant to the application should also be considered in the selection process (e.g., normal boiling points for solvent recovery and environmental properties).

As discussed earlier, the Tanimoto similarity index within the GH-GNN model framework can provide additional insights into the risk or confidence associated with its predictions. For all 7 solvents shown in Tab. 2, the Tanimoto index is relatively high and exceeds the recommended value given in [4] (i.e., 0.35). This suggests a high level of confidence in these predictions. Out of the 349 solvents compared here, only 4 have a Tanimoto metric smaller than 0.35.

CONCLUSIONS

The procedure of solvent pre-selection is central to the efficient exploration of the vast chemical space, enabling efficient use of resources during the early stages of process design. Predictive thermodynamic methods, such as UNIFAC-based models, are typically employed for this purpose. However, their use is limited by the extent to which the involved molecules can be successfully fragmented into functional groups and the availability of the necessary interaction parameters. The accuracy of such methods is also limited when compared to recent alternatives developed by use of deep learning approaches.

A hybrid graph neural network model, the GH-GNN [4], was here extended to accommodate the simultaneous prediction of organic and ionic systems. Three different extension strategies were compared, and a highly accurate GH-GNN model that can simultaneously predict organic and ionic systems was developed. We demonstrate that while the selection of solvents between UNIFAC-IL and GH-GNN overlaps to a great extent, there are also significant differences that might influence the outcomes of the early stages of process design. The higher accuracy of GH-GNN and larger applicability across the chemical space when compared to UNIFAC-IL, makes the GH-GNN model a potentially more reliable alternative for solvent pre-selection.

In the context of sustainable chemistry, we have chosen a relevant case study for the extraction of the monomer caprolactam from the ionic liquid [EMIM][BF4].

Table 2: Top 7 solvents selected by the GH-GNN and UNIFAC-IL. The ranking of the respective solvent for both GH-GNN and UNIFAC-IL is shown. Tanimoto similarity index of the GH-GNN model is also presented.

Solvent	Rank GH-GNN	Rank UNIFAC-IL	Tanimoto index
Diisononyl adipate	1	3	0.88
Eicosane	2	1	1
Hexadecane	3	2	1
Stearic acid	4	4	0.96
Dodecane	5	5	1
Tridecyl alcohol	6	6	1
Dibutyl sebacate	7	7	0.98

The top 7 solvents by both GH-GNN and UNIFAC-IL predictions were given for this case study and the Tanimoto index was close to 1 for most of these, showing a high confidence level. The need for such an extraction process arises in the context of chemical recycling of polyamide 6, highlighting the enormous potential of models like GH-GNN to support chemical engineers in developing more powerful and sustainable separation processes.

ACKNOWLEDGEMENTS

This work was partly supported by the Research Initiative "SmartProSys: Intelligent Process Systems for the Sustainable Production of Chemicals", funded by the Ministry for Science, Energy, Climate Protection and the Environment of the State of Saxony- Anhalt.

REFERENCES

1. Sanchez Medina, E.I. and Sundmacher, K., 2023. Solvent pre-selection for extractive distillation using Gibbs-Helmholtz Graph Neural Networks. In Computer Aided Chemical Engineering (Vol. 52, pp. 2037-2042). Elsevier.
2. Sanchez Medina, E.I., Linke, S., Stoll, M. and Sundmacher, K., 2022. Graph neural networks for the prediction of infinite dilution activity coefficients. Digital Discovery, 1(3), pp.216-225.
3. Rittig, J.G., Hicham, K.B., Schweidtmann, A.M., Dahmen, M. and Mitsos, A., 2023. Graph neural networks for temperature-dependent activity coefficient prediction of solutes in ionic liquids. Computers & Chemical Engineering, 171, p.108153.
4. Sanchez Medina, E.I., Linke, S., Stoll, M. and Sundmacher, K., 2023. Gibbs–Helmholtz graph neural network: capturing the temperature dependency of activity coefficients at infinite dilution. Digital Discovery, 2(3), pp.781-798.
5. Sanchez Medina, E.I., Kunchapu, S. and Sundmacher, K., 2023. Gibbs–Helmholtz graph neural network for the prediction of activity coefficients of polymer solutions at infinite dilution. The Journal of Physical Chemistry A, 127(46), pp.9863-9873.
6. Chen, G., Song, Z., Qi, Z. and Sundmacher, K., 2021. Neural recommender system for the activity coefficient prediction and UNIFAC model extension of ionic liquid-solute systems. AIChE Journal, 67(4), p.e17171.
7. Qin, S., Jiang, S., Li, J., Balaprakash, P., Van Lehn, R.C. and Zavala, V.M., 2023. Capturing molecular interactions in graph neural networks: a case study in multi-component phase equilibrium. Digital Discovery, 2(1), pp.138-151.
8. Minor, A.J., Goldhahn, R., Rihko-Struckmann, L. and Sundmacher, K., 2023. Chemical recycling processes of Nylon 6 to Caprolactam: Review and Techno-Economic assessment. Chemical Engineering Journal, p.145333.
9. Gmehling, J. and Schedemann, A., 2014. Selection of solvents or solvent mixtures for liquid–liquid extraction using predictive thermodynamic models or access to the Dortmund Data Bank. Industrial & Engineering Chemistry Research, 53(45), pp.17794-17805.

© 2025 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See https://creativecommons.org/licenses/by-sa/4.0/