Systems & Control Transactions

Research Article
Foundations of Computer Aided Process Design (FOCAPD 2024)
Breckenridge, Colorado, USA. July 14-18, 2024
Peer Reviewed Conference Proceeding

PSE PRESS

Machine Learning-Aided Process Design for Microwave-Assisted Ammonia Production

Md Abdullah Al Masud, Alazar Araia, Yuxin Wang, Jianli Hu, Yuhe Tian*

West Virginia University, Department of Chemical and Biomedical Engineering, Morgantown, West Virginia, USA
* Corresponding Author: yuhe.tian@mail.wvu.edu.

ABSTRACT

Machine learning (ML) has become a powerful tool to analyze complex relationships between multiple variables and to unravel valuable information from big datasets. However, an open research question lies in how ML can accelerate the design and optimization of processes in the early experimental development stages with limited data. In this work, we investigate the ML-aided process design of a microwave reactor for ammonia production with exceedingly little experimental data. We propose an integrated approach of synthetic minority oversampling technique (SMOTE) regression combined with neural networks to quantitatively design and optimize the microwave reactor. To address the limited data challenge, SMOTE is applied to generate synthetic data based on experimental data at different reaction conditions. Neural network has been demonstrated to effectively capture the nonlinear relationships between input features and target outputs. The softplus activation function is used for a smoother prediction compared to the Rectified Linear Unit activation function. Ammonia concentration is predicted using pressure, temperature, feed flow rate, and feed composition ratio as input variables. For point-wise prediction based on discrete operating conditions, the proposed SMOTE integrated neural network approach outperforms with 96.1% accuracy compared to neural networks (without SMOTE), support vector regression, and linear regression. The multi-variate prediction trends are also validated which are critical for design optimization.

Keywords: Process Design, Process Intensification, Machine Learning, Neural Networks, Ammonia Production

INTRODUCTION

Machine Learning (ML) offers the capability to surpass the constraints of mechanistic modeling by enabling the learning of complex relationships between process variables and the target outputs, offering cost-effective model development, and proving advantageous for optimization [1-2]. To reliably transform the data into valuable predictions, ML aims to acquire rules and patterns from a sufficiently large number of samples [3-4]. However, in certain cases, data availability may be intrinsically limited such as during the early development stage of novel experimental technologies. The scarcity of the datasets may hinder the accuracy of ML predictions. The prediction error typically follows a consistent power-exponential decline as the dataset size increases [5]. To give a more intuitive idea, doubling the number of training samples can result in approximately 20% reduction in prediction error [6]. As such, a key research challenge is how ML can successfully aid process design and optimization at early-stage process developments using sparse experimental data.

To address this challenge, several approaches have been developed in recent research work which can be categorized as: (i) Sampling, encompassing both over and under sampling techniques [7,8], (ii) Cost sensitivity, which involves modifying the cost function to prompt models to prioritize minority samples [9], (iii) Adversarial network generation, which can produce spurious data [10]. However, these methods show inefficacy in handling continuous target output prediction in the regression context which is essential for the data-driven modeling of chemical process systems.

In this work, we present a neural network-based method integrating synthetic minority oversampling technique (SMOTE). The proposed strategy is applied to the design of a microwave-assisted ammonia production process with little experimental data. The remainder of this paper is organized as follows: Section 2 introduces the motivating case study of a microwave-assisted reactor for ammonia production. Section 3 details the SMOTE-integrated neural network method. Section 4 showcases this method for the data-driven modeling and optimization of the microwave reactor. Section 5 presents concluding remarks and ongoing work.

CASE STUDY: MICROWAVE-ASSISTED AMMONIA PRODUCTION

The use of microwaves in chemical synthesis has sparked considerable interest in recent years. Compared with traditional thermal heating, microwave offers the capability of direct and selective volumetric heating which can result in higher reaction selectivity, shorter reaction times, and milder reaction conditions [11,12]. One of the key applications of microwave reactor is ammonia production under ambient/moderate pressure and moderate temperature [13]. This provides a promising technological alternative to the current ammonia synthesis route via the highly energy intensive Haber-Bosch process.

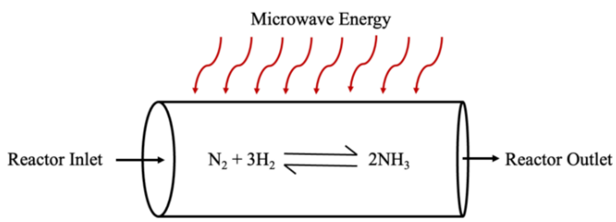

Figure 1. Schematic of a cylindrical microwave reactor.

A schematic of the microwave-assisted ammonia reactor is shown in Fig. 1 which is adapted as the case study of this work [14]. Ru-based catalyst is used which has been demonstrated to exhibit high activity for ammonia synthesis under moderate reaction conditions, particularly when alkali metal Cs is used as a promoter. A total of 46 data points is collected from the microwave reactor at different operating conditions. Table 1 gives examples of such measured data. Our research objective is to predict the optimal reaction conditions for ammonia synthesis based on these experimental data.

Table 1: Experimental data – an indicative list.

T (°C)	F (ml/h·g_{cat})	NH_3%	F (ml/h·g_{cat})	NH_3%
280		0.56		0.27
320		1.57		1.02
340	6000	1.63	15000	1.21
360		1.55		1.25
400		1.28		1.06

THE PROPOSED METHODOLOGY

As depicted in Fig. 2, a neural network approach is developed leveraging SMOTE and softplus activation function. Neural network is employed to capture the nonlinear input-output relationship. SMOTE regression [6,7] is applied to overcome the challenge of limited data availability, which can generate new synthetic data to be used for ML training together with the original experimental data. Soft plus activation function is applied to obtain smooth and differentiable output predictions. In what follows, we briefly introduce the key components of the proposed methodology.

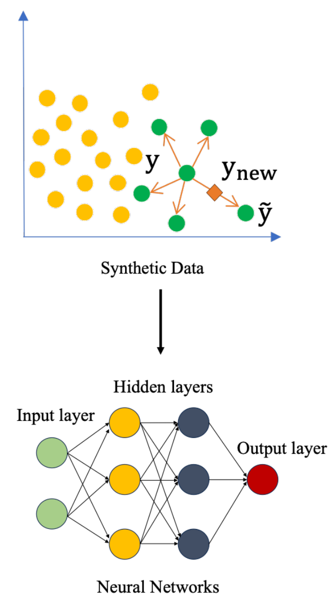

Figure 2. SMOTE integrated Neural Networks.

SMOTE Regression

SMOTE [7] is commonly employed for data simulation to create distribution-dependent neighbor samples for the minority class. Unlike the random oversampling algorithms, SMOTE can dynamically generate the necessary number of samples for the minority class. The use of SMOTE aims to explore new samples that closely resemble the original data distribution. This is instrumental for generating experimentally relevant data with similarities to the existing dataset. In the data generation algorithm of SMOTE, several samples are randomly chosen from the k neighbors for each sample y, in the minority class. A new sample is constructed for each randomly selected neighbor from the original sample, as shown by Eq. 1. The new synthetic data points are then merged with the original experimental data to be fed to neural network training.

$$y_{new} = y + rand(0,1) \times (y - \tilde{y}) \quad (1)$$

Neural Network with Soft plus Activation

Neural network models are comprised of nodes, where each node has the capacity to host several neurons. The neural network architecture used in this work comprises of one input layer, two hidden layers, and one output layer. Each neuron is associated with weights and biases which get updated to minimize the difference between the true labels and predicted labels of the input datasets. The output of a neuron becomes input for the consecutive neurons. The output is expressed by Eq. 2.

$$y^t = softplus(w^t x^t + b^t) \quad (2)$$

where w^t is the weight vector, x^t is the vector for input data, b^t is the bias.

The role of activation function in a neural network architecture is that it transforms the linear combination of summed weights and biases into a nonlinear output. Activation function normalizes the data that imposes a restriction to convert the output of a neuron into a specific bound. The Rectified Linear Unit (ReLU) and Softplus activation functions are compared as shown in Fig. 3. ReLU activation function is a non-differentiable function which may provide non-smooth predictions compared to Softplus, i.e. $f(x) = log\ (1 + e^x)$. A major advantage of using Softplus is that it does not suffer from the "dying ReLU" problem which refers to the situation when some of the neurons become inactive, and they cannot update their weights and biases resulting in non-smooth predictions.

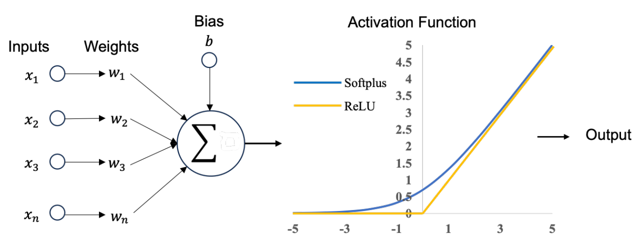

Figure 3: Comparison of activation functions.

MACHINE LEARNING-AIDED MICROWAVE DRIVEN REACTION DESIGN

In this section, we apply the SMOTE-integrated neural network approach to model the microwave-assisted ammonia reaction. Specifically, we will investigate:

a. Point-wise prediction for NH_3 concentration based on discrete operating conditions (i.e., temperature, pressure, inlet H_2 to N_2 ratio, inlet gas flow rate).
b. Continuous prediction to capture the variation trend of NH_3 concentration under varying operating conditions.
c. Reaction design optimization for maximum NH_3 concentration.

Point-wise Prediction

Two hidden layers are used for the neural networks with 128 and 64 nodes, respectively. 40,589 synthetic data are generated via SMOTE. The training and testing data are separated by an 80-20% split. The adam optimizer is applied utilizing the mean squared error as the loss function. The number of epochs is set as 100 with a learning rate of 0.0001. The average discrete prediction accuracy of SMOTE-integrated neural network is 96.10% as shown in Table. 1. This accuracy is superior to other regression methods such as neural networks (without SMOTE, 95.10%), support vector regression (88.70%), and linear regression (86.30%).

Table 2: SMOTE-integrated NN pointwise prediction.

Actual NH_3 Concentration	Reaction Conditions [T(°C), P(psig), R, F (ml/h·g_{cat})]	Prediction Accuracy (%)
0.63	320, 20, 0.50, 6000	99.60
0.56	320, 0, 3, 6000	95.50
1.55	320, 75, 1, 6000	99.80
1.35	320, 80, 2, 6000	92.92
0.60	320, 0, 2, 6000	95.80
0.89	320, 40, 3, 6000	98.68
1.06	320, 80, 3, 6000	95.14
1.57	320, 80, 1, 6000	92.31
…	…	…
		Avg. = 96.10

Continuous Prediction

Herein, the primary goal is to verify that the resulting data-driven model can also capture the continuous variational trend of ammonia concentration versus operating conditions. We first develop a SMOTE-integrated neural network model for two inputs and one output. For example, NH_3 concentration is computed from the two-input one-output data-driven model at varying pressures and inlet H_2 to N_2 ratios while temperature and inlet gas flow rate are kept constant. Two separate models are developed for the prediction of NH_3 concentration, respectively using 11,340 and 4,180 synthetic data generated by SMOTE. The total number of synthetic data is reduced to avoid overfitting. The prediction results are validated against the original experimental data as shown in Fig. 4, in which the markers represent the original experimental data, and the solid lines depict the continuous prediction using the above trained SMOTE-integrated NN model.

On this basis, we proceed to build a single data-driven model considering all the four input variables (i.e., temperature, pressure, H_2 to N_2 ratio, and feed flow rate).

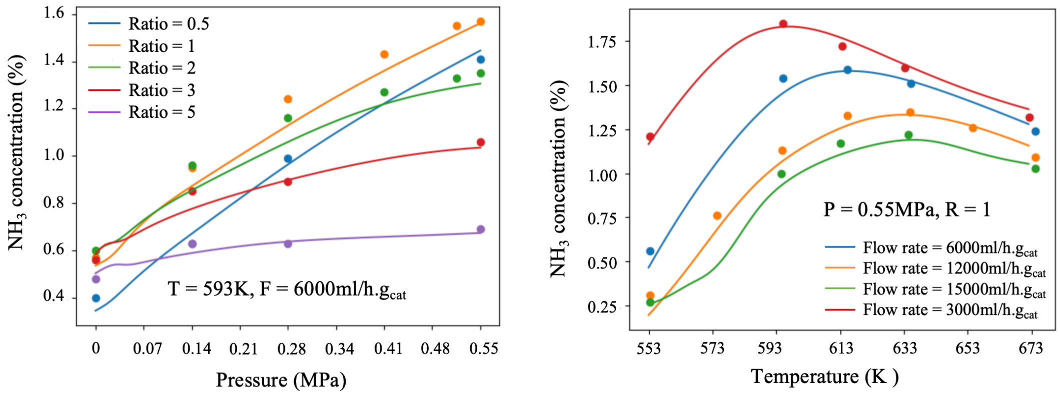

Figure 4: Continuous prediction with two-input one-output data-driven model. (Markers: Original experimental data, Solid lines: SMOTE integrated NN prediction)

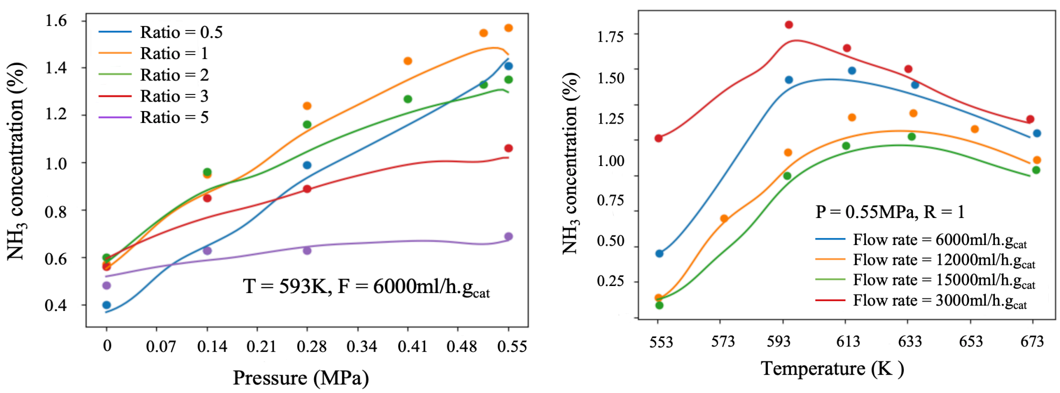

Figure 5: Continuous prediction with four-input one-output data-driven model.

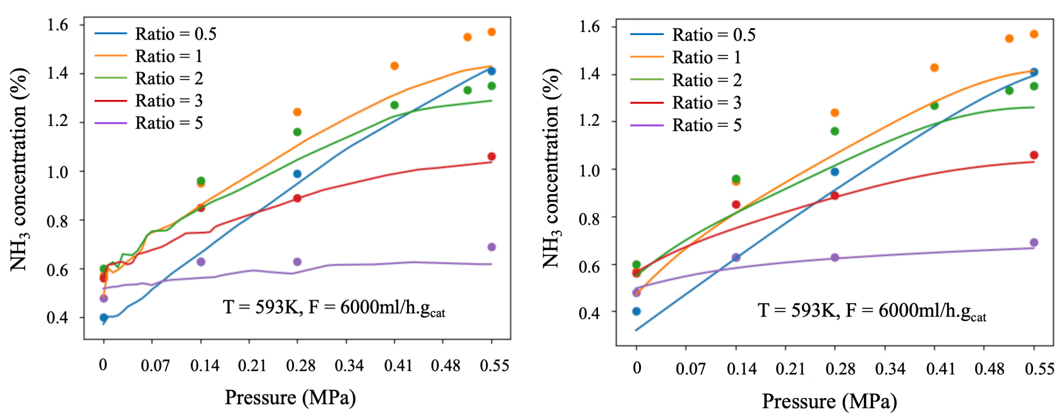

Figure 6: Role of activation function. (Left) ReLU. (Right) Soft plus.

As the input dimensions increase, the input-output relationship becomes more nonlinear and complex. Namely, the resulting data-driven model should simultaneously capture the nonlinear relationship between ammonia concentration versus temperature, pressure, inlet H_2 to N_2 ratio, and inlet gas flow rate. 23,000 synthetic data generated via SMOTE are used to train the four-input one-output neural network. As illustrated in Fig. 5, the continuous prediction using a higher dimensional data-driven model can also effectively capture the variational trend while featuring larger deviations than the two-input one-output models. The role of activation functions is also investigated as shown in Fig. 6. It can be noted that the use of ReLU activation function may result in non-smooth predictions compared with that of Softplus activation function. This is consistent with the observations reported in open literature [15].

Comparison with SVR and NN

The continuous predictions from our proposed methodology were compared to the support vector regression (SVR) and NN (without SMOTE). For SVR, Radial Basis Function (RBF) is identified as the best kernel function for our datasets. As illustrated in Fig. 7, SVR identifies a pseudo-linear multi-variate relationship instead of capturing the correct nonlinear trends. The predictions using NN (without SMOTE) are depicted in Fig. 8. Even if NN offers a point-wise prediction accuracy of 95.10%, it introduces excessive nonlinear directional changes in the continuous prediction. In other words, NN cannot effectively capture the continuous variational trend. If this data-driven model is applied for design optimization, it may fail to identify the correct direction toward optimality.

Design Optimization

Finally, we performed the design optimization to determine the optimum reaction conditions. A mesh grid of data points is created and fed to the afore-trained SMOTE-integrated neural networks (Fig. 9). The ammonia concentration for each of the input data is obtained from the data-driven model prediction and the index of maximum ammonia concentration is located. The optimum temperature, pressure, H_2 to N_2 ratio, and inlet gas flow rate are identified as 324.37°C, 80psig, 1, and 3000ml/h·g_{cat} respectively. The optimum operating conditions obtained from experimental data are at the temperature of 320°C, pressure of 80psig, H_2 to N_2 ratio of 1, and inlet gas flow rate of 3000ml/h·g_{cat} which well justifies the validity of data-driven model prediction.

CONCLUSION

In summary, this work has developed a machine learning-aided method to design and optimize microwave-assisted ammonia reaction conditions with little experimental data. The efficacy of the proposed approach is demonstrated on the point-wise and continuous prediction of ammonia concentration under varying reaction temperature, pressure, H_2 to N_2 ratio, and feed flow rates. Ongoing work is investigating the systems-level analysis leveraging this data-driven microwave reactor model.

ACKNOWLEDGEMENT

This research project acknowledges the funding resource of West Virginia University and WV Research Challenge Grant (RCG23-009): Metal-embedded carbon-based catalytic membranes for co-production of ammonia and ethylene.

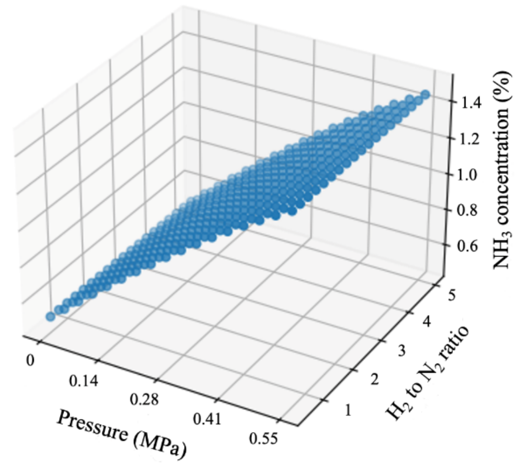

Figure 7: Continuous prediction using SVR.

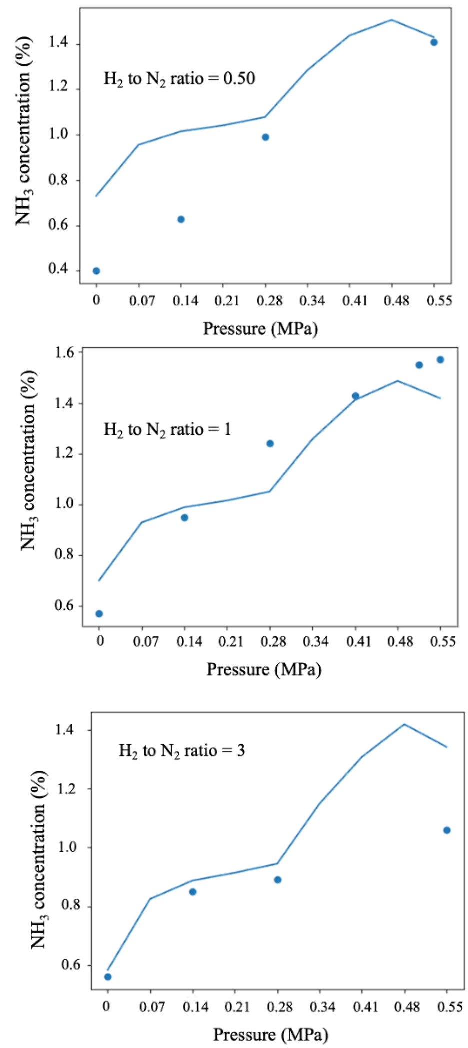

Figure 8: Continuous prediction using NN without SMOTE. (Markers: Original experimental data, Solid lines: NN prediction

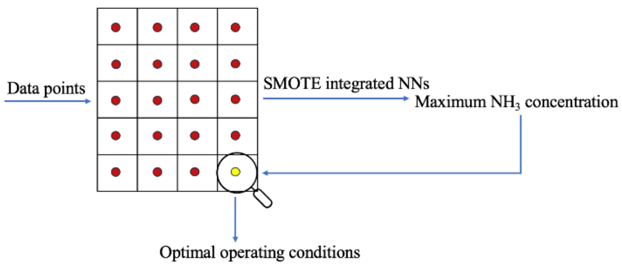

Figure 9: Illustration of design optimization.

REFERENCES

1. Schweidtmann, A. M., Esche, E., Fischer, A., Kloft, M., Repke, J. U., Sager, S., & Mitsos, A. (2021). Machine learning in chemical engineering: A perspective. *Chemie Ingenieur Technik*, *93*(12), 2029-2039.
2. Lee, J. H., Shin, J., & Realff, M. J. (2018). Machine learning: Overview of the recent progresses and implications for the process systems engineering field. *Computers & Chemical Engineering*, *114*, 111-121.
3. Golkarnarenji, G., Naebe, M., Badii, K., Milani, A. S., Jazar, R. N., & Khayyam, H. (2019). A machine learning case study with limited data for prediction of carbon fiber mechanical properties. *Computers in Industry*, *105*, 123-132.
4. Zhang, Z., Schott, J. A., Liu, M., Chen, H., Lu, X., Sumpter, B. G., & Dai, S. (2019). Prediction of carbon dioxide adsorption via deep learning. *Angewandte Chemie*, *131*(1), 265-269.
5. Faber, F. A., Lindmaa, A., Von Lilienfeld, O. A., & Armiento, R. (2016). Machine learning energies of 2 million elpasolite (ABC_2D_6) crystals. *Physical review letters*, *117*(13), 135502.
6. Liu, X., Xu, Y., Li, J., Ong, X., Ibrahim, S. A., Buonassisi, T., & Wang, X. (2021). A robust low data solution: dimension prediction of semiconductor nanorods. *Computers & Chemical Engineering*, *150*, 107315.
7. SMOTE-RSB*: a hybrid preprocessing approach based on oversampling and undersampling for high imbalanced data-sets using smote and rough sets theory. *Knowledge and information systems*, *33*, 245-265.
8. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321-357.
9. Elkan, C. (2001, August). The foundations of cost-sensitive learning. In *International joint conference on artificial intelligence* (Vol. 17, No. 1, pp. 973-978). Lawrence Erlbaum Associates Ltd.
10. Antoniou, A., Storkey, A., & Edwards, H. (2017). Data augmentation generative adversarial networks. *arXiv preprint arXiv:1711.04340*.
11. Horikoshi, S., & Serpone, N. (2015). General Introduction to Microwave Chemistry. *Microwaves in Catalysis: Methodology and Applications*, 1-28.
12. Hunt, J., Ferrari, A., Lita, A., Crosswhite, M., Ashley, B., & Stiegman, A. E. (2013). Microwave-specific enhancement of the carbon–carbon dioxide (Boudouard) reaction. *The Journal of Physical Chemistry C*, *117*(51), 26871-26880.
13. Wang, Y., Khan, T. S., Wildfire, C., Shekhawat, D., & Hu, J. (2021). Microwave-enhanced catalytic ammonia synthesis under moderate pressure and temperature. *Catalysis Communications*, *159*, 106344.
14. Hu, J., Wildfire, C., Stiegman, A. E., Dagle, R. A., Shekhawat, D., Abdelsayed, V., & Wang, Y. (2020). Microwave-driven heterogeneous catalysis for activation of dinitrogen to ammonia under atmospheric pressure. *Chemical Engineering Journal*, *397*, 125388.
15. Rittig, J. G., Felton, K. C., Lapkin, A. A., & Mitsos, A. (2023). Gibbs–Duhem-informed neural networks for binary activity coefficient prediction. *Digital Discovery*. 2, 1752-1767.