Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Predicting The Tensile And Flexural Strength Of Fire-Retardant Epoxy Materials Using A Data-Driven Approach

Pooya Jafari [1], Pingan Song [2]

[1] University of Southern Queensland, Centre for Future Materials, Springfield, 4300, Australia
[2] the University of Southern Queensland, School of Agriculture and Environmental Science
[1] pooya.jafari@unisq.edu.au; [2] pingan.song@unisq.edu.au

Modeling and predicting the mechanical properties of materials, particularly polymers, is a longstanding challenge. The mechanical properties of polymer materials can be influenced by various factors, making it essential to maintain their integrity across different applications. Fire retardant polymeric materials, in particular, may experience changes in their mechanical properties due to the addition of fire retardants designed to reduce flammability. To address this, we have chosen to adopt a data-driven approach to predict the tensile and flexural strength of these materials using artificial intelligence. This innovative and promising method is increasingly employed by researchers to tackle a wide range of scientific problems.

Keywords— In Fire retardant- Machine learning- Mechanical properties- Modelling, Artificial Inteligence“,”

I. INTRODUCTION

Since their initial discovery, polymeric materials have become immensely prevalent in modern society. These versatile substances have integrated into nearly every facet of contemporary life, being extensively utilized across a wide array of industries and sectors [1]. Recent advancements in polymeric materials offer numerous benefits to modern communities. These materials are extensively used in various applications, including manufacturing, construction, healthcare, electronics, consumer goods, transportation, and building [2]. However, polymeric materials can present considerable safety hazards in applications requiring high flame resistance due to their inherent flammability [3]. New heat-resistant polymer materials need to exhibit excellent thermal stability and significant processability. However, achieving high thermal resistance in polymers can result in undesirable, weaker processing properties [4]. Developing flame-retardant polymers to meet high standards has consistently been challenging because traditional methods rely on experiential intuition and trial-and-error screenings, which are time-consuming [5]. Based on the structure of the training data, machine learning is typically separated into three sub-groups: unsupervised, supervised, and semi-supervised [6]. In supervised machine learning (ML), data are

categorized, and predictions are generated using correctly labelled datasets. In contrast, unsupervised ML, which does not utilize labelled data, focuses on identifying relationships within datasets. Although supervised ML demands more resources due to the need for labelled data, it can yield exceptional results when such training data are accessible. Popular algorithms for designing and developing advanced materials include decision trees, support vector machines (SVM), artificial neural networks, ensemble learning, and clustering [7].

Recently, there has been increasing interest in using Machine Learning (ML) models to explore new materials. The domain of artificial intelligence, especially machine learning (ML), has been widely applied across numerous scientific research fields [8]. This approach has become prominent for its practical benefits in improving material property design, thanks to advancements in computing power and related algorithms [9]. Large datasets, feature engineering, and ML regression techniques can all be used to forecast material properties for practical applications and efficient production. Researchers can save time and effort in their experimentation by training on preexisting data. These methods aid in the discovery and optimization of functional materials for use in thermoelectric, solar, catalytic, and optical applications, among other domains [10]. In polymer science, Machine Learning has become increasingly significant in recent years and continues to evolve. For instance, Zhu et al. [11] developed an ML algorithm specifically to rationally design polymer nanocomposites. Wei et al. [12] systematically investigated a Machine Learning algorithm to classify various states of polymer configurations.

In this paper, we decided to use a machine learning model in order to investigate the tensile and flexural strength of fire-retardant epoxy resins. The conventional method for modeling the mechanical properties of materials involves numerical approaches like molecular dynamics simulations. However, these methods are highly time-consuming and expensive, requiring high-performance computing resources. To address these challenges, we propose a data-driven approach, which significantly reduces both time and cost. By leveraging machine learning algorithms and datasets, this approach can efficiently predict material properties, providing a faster and more cost-effective alternative to traditional numerical simulations. Furthermore, when new data become available, the data-driven model may be continuously enhanced and updated to increase its forecast accuracy and dependability. The flammability properties such limiting oxygen index (LOI), peak heat release rate (PHRR), total heat release (THR), time to ignition (TTI), and vertical combustion test (UL94) level have been investigated by Chen et al. [13] before for epoxy resin composites. Based on the same approach, we decided to investigate their tensile and flexural strength.

II. METHODS AND RESULTS

A. Method

Chen et al. [13] gathered the molecular structures of 315 fire retardant molecules and changed them into a Simplified Molecular Input Line Entry System (SMILES) format which can be used by computer programs [14]. A molecule is depicted by SMILES as a line of text that contains connectivity, bond kinds, and atomic symbols [15]. An example has been presented in Fig. 1 for clarification. Using these molecules provided in SMILES format, various computer programs can calculate their chemical properties, enabling the creation of a database

for training a machine learning model. They used alvaDesc [16] in order to calculate 284 molecular descriptors such as the molecular weight, the average molecular weight, the number of Oxygen, and more. Moreover, they added another column to their database for the addition amount of fire retardant in epoxy resin composites. Based on these variables, they created a database and decided to predict the flammability properties of these materials as stated before. This inspired us to predict the mechanical properties of these materials based on the database they provided. Thus, we gathered the data related to the tensile and flexural strength of the fire-retardant epoxy composites based on the same input variables. However, we realized that the tensile and flexural strength reported in different papers for fire retardant epoxy composites is different for pure epoxy which has no fire-retardant addition amount. This might decrease the efficiency and accuracy of the machine learning model which made us to calculate a delta formula for the tensile strength and flexural strength of each molecule. This formula is depicted as follows:

$$\Delta\sigma\ (\%) = \frac{\sigma_{\text{fire retadant composite}}(\text{MPA}) - \sigma_{\text{pure epoxy}}(\text{MPA})}{\sigma_{\text{pure epoxy}}(\text{MPA})}$$

$$\Delta\sigma_{\text{f}}\ (\%) = \frac{\sigma_{\text{f(fire retardant composite)}}(\text{MPA}) - \sigma_{\text{f(pure epoxy}}(\text{MPA})}{\sigma_{\text{f(pure epoxy)}}(\text{MPA})} \quad (2)$$

Based on the formula above, the increase percentage is measured, which can be more accurate considering the explanations provided earlier. Similar to the database in Chen et al. [13] publication, we created a database including 160 samples and 286 input variables with the target variable $\Delta\sigma$ (%) to be used for training a machine learning algorithm. Moreover, we created another dataset with 138 samples and 286 input variables for the delta flexural strength $\Delta\sigma_{\text{f}}$ (%).

To train the model using the provided database, we divided the dataset into two subsets: training and test sets. The training set is used to train the model and the test set is used to evaluate the performance and accuracy of the model. This is recommended by those experts in the artificial intelligence field for a better training and better understanding of the model's performance [17]. It is usually suggested to use 25% of the dataset for the test and the rest for training. However, this percentage can be varying according to the number of datapoints [18]. For instance, in cases with a low number of data points, such as ours, it is recommended to use 15% of the dataset for testing and the remaining 85% for training [18]. This allows the model to be trained with more data enabling it to have better performance and accuracy.

The dataset needs to be feature scaled before training to adjust the range of the raw data values. This process helps improve the model's prediction efficiency. After feature scaling, a feature selection technique needed to find and choose the most important input variables for predicting the target variable. We used the Recursive Feature Elimination (RFE) for this matter as it employes an estimator, an algorithm, to be trained on the initial set of features. In our study, we used Ridge algorithm [19] as it is one of those fundamental algorithms that every skilled data scientist should be able to use. The importance of each feature is assessed, and those features that have less impact on the target variable are removed. By removing the

less effective features, the model can consider the most significant ones leading to a better performance and accuracy. After this step, the main model training would begin, and the performance would be evaluated. There are different algorithms to choose for the model such LASSO [20], Ridge, the Artificial neural network [21], and more. We used a Multi-Layer Perceptron (MLP), a type of artificial neural network which can be used for regression problems. This algorithm consists of multiple layers of neurons dividing into three layers: input, hidden and output layer. The firs layer receives the input features where each neuron is specified for one feature. Next, one or more layers of neurons (hidden layers) apply transformation and some activation functions on the variables. In the end, an output layer produces the results based on what happened in pervious layers. A schematic of Multi-Layer Perceptron (MLP) is presented in Fig. 2.

After finishing the training, the performance of the model needs to be evaluated. In fact, it is important to know how close predictions are to actual results. There are different statical values in order to assess the performance and accuracy of the model such as root-mean-square error (RMSE), mean absolute error (MAE), and more. In our study, we used the mean square error (MSE) or root-mean-square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2). These statistical values are explained as follows:

$$R^2 = \frac{\sum_{i=1}^{n}(\widehat{y_i} - \overline{y})^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{3}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\widehat{y_i} - y_i)^2}{n}} \tag{4}$$

$$MAE = \frac{\sum_{i=1}^{n}|\widehat{y_i} - y_i|}{n} \tag{5}$$

$$MSE = \frac{\sum_{i=1}^{n}(\widehat{y_i} - y_i)^2}{n} \tag{6}$$

In the equations above, n shows the total number of data points, $\widehat{y_i}$ is the predicted value for each point, y_i is the experimental value for each point, and $\overline{y}$ is the mean value of all experimental cases [1].

B. Results

After training the model on the provided dataset, we assessed its performance using the statistical metrics discussed earlier. Fig. 3 illustrates the model's effectiveness on both the training and test datasets. The coefficient of determination (R^2) is 0.927 for the training set and 0.784 for the test set, indicating a high level of accuracy in predicting the target variable.

To further confirm the model's accuracy, additional statistical measures were evaluated. The Mean Squared Error (MSE) is 119.5, the Mean Absolute Error (MAE) is 7.68, and the Root Mean Squared Error (RMSE) is 10.93. The MAE represents the average absolute difference between the actual and predicted values, which in this case is relatively low at 7.68.

Despite the relatively high MSE, the low MAE suggests that while most predictions are

accurate, there are a few with significantly larger errors. The MSE, being sensitive to larger deviations due to the squaring of errors, underscores the presence of these outliers or substantial prediction errors that are less apparent in the MAE. This discrepancy indicates that while the model performs well on average, there are specific instances with higher prediction errors that could be further investigated to improve overall model robustness.

To further validate the accuracy of the model, we compared its predictions against actual values from a research paper that was not part of the original dataset. This external validation helps to ensure that the model's performance is not limited to the specific data it was trained on and provides an additional layer of verification by testing the model's predictions against independent, real-world data. By doing so, we can assess the model's generalizability and robustness beyond the scope of the training and testing datasets. Table I provides a full summary of the validation's findings. The model's predictions and the actual data are shown in this table, showing how well the forecasts match the true values. The strong correlation highlights the dependability and efficiency of the approach in predicting delta tensile strength. This kind of validation demonstrates that the model not only functions well overall, but also consistently makes accurate predictions. This consistency is necessary to guarantee the model's applicability in scenarios where precise delta tensile strength prediction is required. Despite the model's relatively high-performance following training, it's important to recognize that it may not reliably predict the delta tensile strength for all fire-retardant molecules. The model's generalization capability could be limited due to the relatively small size of the training dataset. Additionally, potential inaccuracies in the reported results could adversely affect the model's overall performance. Even though we have employed various strategies such as feature scaling, feature selection, and using delta tensile strength as the target variable, these measures might not be sufficient to ensure robust generalization. The limited dataset size and possible errors in the data may restrict the model's ability to accurately generalize to new, unseen molecules. Therefore, while the model shows promising results, further efforts may be necessary to enhance its reliability and predictive accuracy across a broader range of fire-retardant compounds. This might include expanding the dataset, refining the feature selection process, or exploring advanced modeling techniques to improve generalization and reduce prediction errors.

Regarding the delta flexural strength, we assessed its performance using the noted statistical metrics. Fig. 4 depicts the model's efficiency on both the training and test datasets. The coefficient of determination (R^2) is 0.921 for the training set and 0.912 for the test set, proving a very high level of accuracy in predicting the delta flexural strength. To further validate the model's accuracy, additional statistical measures were assessed. The Mean Squared Error (MSE) is 59.4, the Mean Absolute Error (MAE) is 5.38, and the Root Mean Squared Error (RMSE) is 7.7. The MAE, which indicates the average absolute difference between the actual and predicted values, is relatively low at 5.38 in this instance. All the statistical metrics indicate that the model performs with high accuracy in predicting the delta flexural strength. In order to assess the generalization capacity of the model for predicting this target variable, we validated the predicted results of the model with actual data of the same research paper which we used in order to validate the delta tensile strength. The results of this study have been gathered in the Table II which shows the excellent performance of the model in predicting the delta flexural strength (Mpa).

It is important to recognize that while the statistical metrics suggest the model performs

efficiently in predicting delta tensile strength, its generalization capability may be limited due to the small size of the training dataset. Furthermore, potential errors in the reported data within our dataset could also undermine the model's ability to generalize. These variables emphasise that in order to improve the model's overall predictive power and dependability, a larger, more precise dataset is required. Further steps to address these problems and enhance model performance could include the use of robust error handling and cross-validation.

III. CONCLUSION

In this paper, we presented a machine learning approach in order to predict the mechanical properties of fire-retardant epoxy composites. The addition of fire retardant into epoxy resin for extinguishing the flammability might lead to mitigating the mechanical properties of the epoxy which might be a disadvantage for its usage in different applications under different loading conditions. As stated before, the traditional

modelling approaches, such as the molecular dynamic can be time and time consuming. We indicated that our model can predict the mechanical properties such as the tensile and flexural strength with a relatively reasonable accuracy. However, our approach cannot be reliable completely since its hugely dependent on the existing findings which might have potential error in their reporting. Moreover, the low number of our samples might not be able to cover the whole reality of mechanical properties of flame-retardant epoxy composites which can reduce the generalization capacity of our developed model.

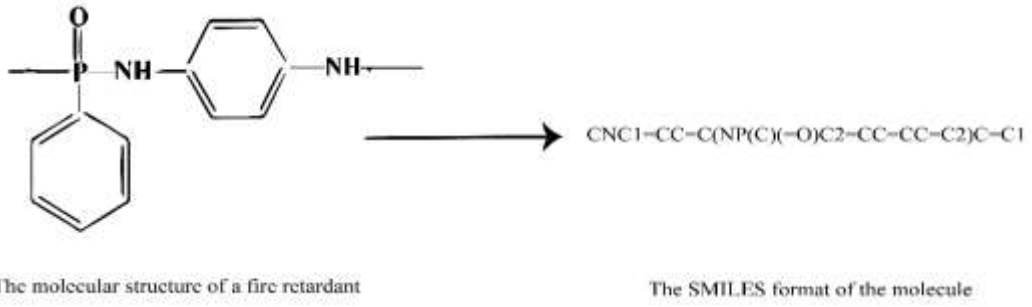

Fig. 1. A schematic of the SMILES format of a fire-retardant molecule

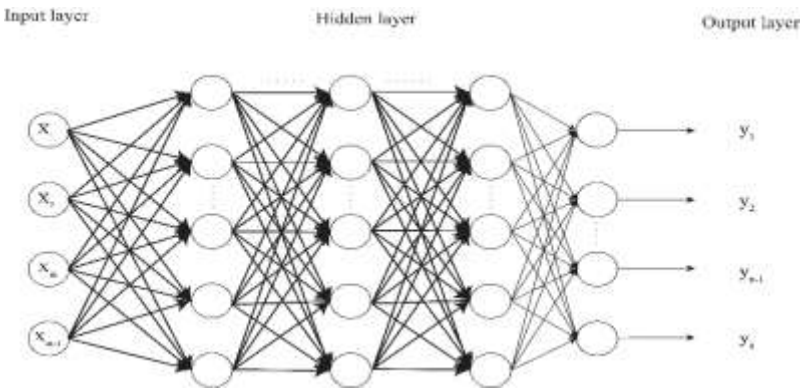

Fig. 2. A diagram of Multi-Layer Perceptron (MLP)

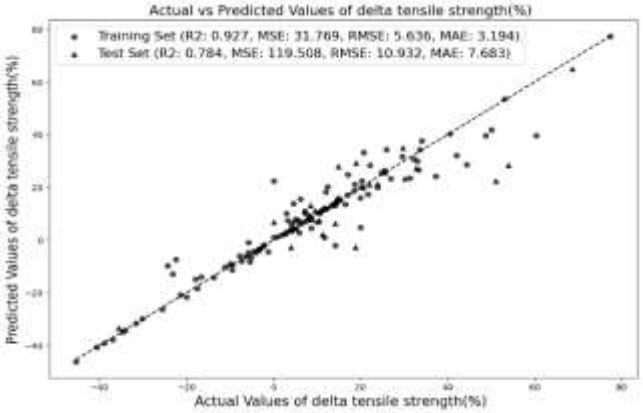

Fig. 3. Statistical performance of the model for predicting the delta tensile strength Δσ (%)

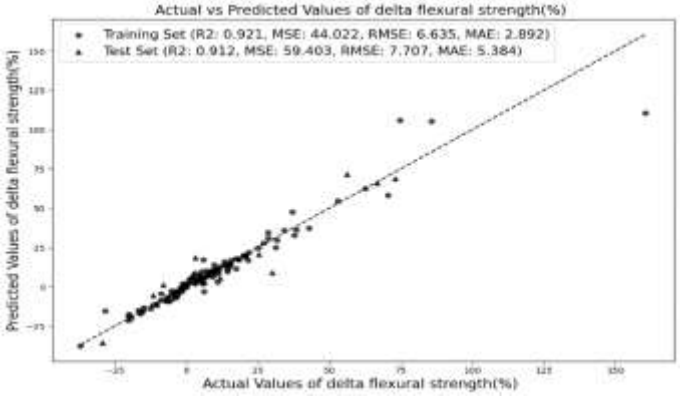

Fig. 4. Statistical performance of the model for predicting the delta flexural strength $\Delta\sigma_f$ (%)

Table I. A comparison study for the delta tensile strength Δσ (%)

Addition amount (wt%)	Li et al. [22]	Present
6	26.92%	27.76%
8	30.76%	32.77%
10	33.04%	37.17%
12	37.06%	40.65%
14	42.83%	43.24%
16	43.7%	44.01%
18	45.8%	43.31%

Table II. A comparison study for the flexural strength $\Delta\sigma_f$ (%))

Addition amount (wt%)	Li et al. [22]	Present
6	12.8%	12.14%
8	15.24%	14.04%
10	15.93%	16.05%
12	17.39%	17.51%
14	20.43%	19.18%
16	21.79%	20.9%
18	22.97%	21.68%

APPENDIX

Appendixes, if needed, appear before the acknowledgment.

ACKNOWLEDGMENT

Instead, write “F. A. Pooya Jafari thanks the Australian Research Council (ARC), the sponsor and financial suuporter of this work.

REFERENCES

[1] P. Jafari, R. Zhang, S. Huo, Q. Wang, J. Yong, M. Hong, R. Deo, H. Wang, P. Song, “Machine learning for expediting next-generation of fire-retardant polymer composites,” Composites Communications, vol.45, pp. 101806, 2024.

[2] R. Sun, Y. A. Elabd, “Synthesis and High Alkaline Chemical Stability of Polyionic Liquids with Methylpyrrolidinium, Methylpiperidinium, Methylazepanium, Methylazocanium, and Methylazonanium Cations,” ACS Macro Letters, vol.8, no.5, pp. 540-545, 2019.

[3] S. Liu, V. S. Chevali, Z. Xu, D. Hui, H. Wang, "A review of extending performance of epoxy resins using carbon nanomaterials," Composites Part B: Engineering, vol.136, pp. 197-214, 2018.
[4] Y. Guo, Z. Chen, Y. Yu, T. Chen, Q. Zhang, C. Li, J. Jiang, "A new-type terephthalonitrile derivative flame retardant of bi-DOPO compound with hydroxyl and amino groups on epoxy resin," Journal of Applied Polymer Science, vol.139, no.21, pp. 52209, 2022.
[5] S. Zhang, F. Chu, Z. Xu, Y. Zhou, W. Hu, Y. Hu, "Interfacial flame retardant unsaturated polyester composites with simultaneously improved fire safety and mechanical properties," Chemical Engineering Journal, vol.426, pp. 131313, 2021.
[6] M. Z. Naser, "Mechanistically Informed Machine Learning and Artificial Intelligence in Fire Engineering and Sciences," Fire Technology, vol.57, pp. 2021.
[7] A. R. Katritzky, M. Kuanar, S. Slavov, C. D. Hall, M. Karelson, I. Kahn, D. A. Dobchev, "Quantitative Correlation of Physical and Chemical Properties with Chemical Structure: Utility for Prediction," Chemical Reviews, vol.110, no.10, pp. 5714-5789, 2010.
[8] A. Mannodi-Kanakkithodi, A. Chandrasekaran, C. Kim, T. D. Huan, G. Pilania, V. Botu, R. Ramprasad, "Scoping the polymer genome: A roadmap for rational polymer dielectrics design and beyond," Materials Today, vol.21, no.7, pp. 785-796, 2018.
[9] J. Cai, X. Chu, K. Xu, H. Li, J. Wei, "Machine learning-driven new material discovery," Nanoscale Advances, vol.2, no.8, pp. 3115-3130, 2020.
[10] G. H. Gu, J. Noh, I. Kim, Y. Jung, "Machine learning for renewable energy materials," Journal of Materials Chemistry A, vol.7, no.29, pp. 17096-17117, 2019.
[11] M.-X. Zhu, Q.-C. Yu, H.-G. Song, T.-X. Chen, J.-M. Chen, "Rational Design of High-Energy-Density Polymer Composites by Machine Learning Approach," ACS Applied Energy Materials, vol.4, no.2, pp. 1449-1458, 2021.
[12] Q. Wei, R. G. Melko, J. Z. Y. Chen, "Identifying polymer states by machine learning," Physical Review E, vol.95, no.3, pp. 032504, 2017.
[13] Z. Chen, B. Yang, N. Song, Y. Liu, F. Rong, X. Zhang, T. Chen, Q. Zhang, J. Jiang, T. Chen, Y. Yu, L. X. Liu, "Machine learning-enabled rational design of organic flame retardants for enhanced fire safety of epoxy resin composites," Composites Communications, vol.44, pp. 101756, 2023.
[14] J. Arús-Pous, S. V. Johansson, O. Prykhodko, E. J. Bjerrum, C. Tyrchan, J.-L. Reymond, H. Chen, O. Engkvist, "Randomized SMILES strings improve the quality of molecular generative models," Journal of Cheminformatics, vol.11, no.1, pp. 71, 2019.
[15] K. Handa, M. C. Thomas, M. Kageyama, T. Iijima, A. Bender, "On the difficulty of validating molecular generative models realistically: a case study on public and proprietary data," Journal of Cheminformatics, vol.15, no.1, pp. 112, 2023.
[16] A. Mauri, "alvaDesc: A Tool to Calculate and Analyze Molecular Descriptors and Fingerprints," in Ecotoxicological QSARs, K. Roy Ed. New York, NY: Springer US, 2020, pp. 801-820.
[17] T. Toyao, Z. Maeno, S. Takakusagi, T. Kamachi, I. Takigawa, K.-i. Shimizu, "Machine Learning for Catalysis Informatics: Recent Applications and Prospects," ACS Catalysis, vol.10, no.3, pp. 2260-2297, 2020.
[18] Y. Xu, R. Goodacre, "On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning," Journal of Analysis and Testing, vol.2, no.3, pp. 249-262, 2018.
[19] T. Hastie, "Ridge Regularization: An Essential Concept in Data Science," Technometrics, vol.62, no.4, pp. 426-433, 2020.

[20]J. Ranstam, J. A. Cook, “LASSO regression,” British Journal of Surgery, vol.105, no.10, pp. 1348-1348, 2018.
[21]E. Guresen, G. Kayakutlu, “Definition of artificial neural networks with comparison to other networks,” Procedia Computer Science, vol.3, pp. 426-433, 2011.
[22]Y. Li, H. Zheng, M. Xu, B. Li, T. Lai, “Synthesis of a novel phosphonate flame retardant and its application in epoxy resins,” Journal of Applied Polymer Science, vol.132, no.45, pp. 2015.

First A. Pooya Jafari is a distinguished researcher currently pursuing his Ph.D. in Materials Engineering at the University of Southern Queensland. His thesis, titled "Data-Driven Machine Learning for The Creation of Next-Generation Fire-Retardant Epoxy Resins," is supported by prestigious scholarships, including the USQ318 UniSQ Externally Funded Research Scholarship (ARC Research) and the USQ198 UniSQ International Fees Research Scholarship. With a solid academic foundation, Pooya earned his M.Sc. in Mechanical Engineering (Solid Mechanics) from Shahrekord University, where he ranked 5th among his peers, and his B.Sc. in Mechanical Engineering from the Islamic Azad University (Najaf Abad Branch).

His research interests span a wide range of topics within materials engineering and mechanical behavior, with a particular focus on polymer-based composites, fire retardants, and nanocomposites. His expertise is evident in his numerous publications in esteemed journals, covering advanced research and modeling techniques, such as finite element analysis and machine learning algorithms for predicting mechanical properties. His contributions to the field include groundbreaking work on the vibration analysis of functionally graded graphene platelet reinforced composites and the development of novel fire-retardant polymer composites.

In addition to his academic achievements, Pooya has garnered significant industrial experience. He served as a designer and producer at Solar Energy Spadan Esfahan, where he optimized metal structures for solar panels, and as a production line supervisor at Farpoood Company Esfahan. His practical skills are complemented by a suite of technical proficiencies, including software such as MATLAB, CATIA, and ABAQUS. Beyond his professional endeavors, Pooya is an active English tutor and basketball coach, reflecting his diverse talents and commitment to both intellectual and personal growth.

Second B. Dr Pingan Song is currently Full Professor and Associate Head of School (Research) within the School of Agriculture and Environmental Science. He has been named on the 2023 Clarivate's Highly Cited Researchers list. His research interests include fire retardants and polymers/ polymer composites as well as their structure-property correlations, which aims to solve real-world problems in the areas of electronics, energy, transport, automotive, building and constructions.
He received his PhD degree in Chemistry from Zhejiang University in 2009. He joined University of Southern Queensland (UniSQ) as Senior Research Fellow in 2016, was awarded ARC Future Fellow in 2019, and then promoted to Associate Professor (2020) and Full Professor (2021). Before joining UniSQ, he started his academic career at Zhejiang A & F

University, China as Lecturer (2009), and was promoted to Associate Professor (2011) and Full Professor (2016).

As a lead Chief Investigator (CI) or co-CI, he has successfully secured more than 30 grants including ARC Future Fellowship, ARC DP, ARC LP, ARC Research Hub, and ARC LIFE as well as two National Natural Science Foundation of China (NSFC equivalent to ARC) projects, with a total research funding of over AU$11 million from multiple funding bodies in the last two decades.

Dr Song has co-authored 290+ peer-reviewed publications in high-tier international journals, such as Science, Adv. Mater., Matter, ACS Nano, Adv. Funct. Mater., and Macromolecules. His publications have received 20,000+ times of citations, with an H-index of 87 (Google Scholar). His works have also been highlighted by multiple media outlets, including Science, Chemistry World and MIT Technical Review, and he has also been interviewed by the ABC News and Toowoomba Seven News. On top of that, one of his works has been selected for ARC Annual Report 2021-22.

He has filed over 20 patents, and delivered over 20 invited talks, including keynotes.

He is currently serving as Associate Editor for Sustainable Materials and Technologies (Elsevier, JCR Q1, IF: 9.6), and Frontiers in Materials (Frontiers, JCR Q2, IF: 3.985); Assistant Editor for Composites Part B: Engineering (Elsevier, JCR Q1, IF: 13.1) (handling submissions on fire retardants and polymer composites), and Journal of Materials Science & Technology (Elsevier, JCR Q1, IF: 9.8); and Editorial Advisory Board Member for ACS Sustainable Chemistry & Engineering (American Chemical Society, JCR Q1, IF:8.4).