The Institute for Integrated Energy Systems (IESVic) works on strategic clean technologies, electrification and system integration, built environment, energy-economy-policy modeling, and integrated planning for water-energy-land systems. IESVic provides leadership at the University of Victoria in the study of critical energy issues, human dimensions of energy, education and training, and works closely with industry, not-for-profits, and government.

The IESVic Energy Briefs Series shares research and practice on the development of sustainable energy systems that are reliable, cost-effective and socially acceptable.

Citation: Murphy, L.; Crawford, C. (2025) A divide in focus: Machine learning for energy storage system modelling in academia, a comparison to industry. *IESVic Energy Brief*

Published: August 18, 2025

ISSN 2818-159X (Print)
ISSN 2818-1603 (Online)

A divide in focus: Machine learning for energy storage system modelling in academia, a comparison to industry

Lucas Murphy [1]*, Dr. Curran Crawford[1]
[1] IESVic, University of Victoria, BC, Canada.
* Correspondence: lucasmurphy@uvic.ca

Key messages

- Differing resources and focuses divide machine learning advancement for energy storage between industry and academia
- Pre-trained models and transparency will facilitate the future of machine learning
- Government policies and incentives can help industry and academia collaborate to benefit energy storage modelling as a whole

Importance: The emergence of machine learning in energy storage system modelling

Energy storage system (ESS) operators require models for safe and reliable system functioning. Lithium-ion batteries are one of the most important ESS technologies, and have numerous modelling needs related to their maintenance, control, efficiency, and safety. Data-driven modelling is proving to be an effective method for managing electrochemical systems, as made possible by using existing data acquisition hardware and sensors. Machine learning (ML) provides valuable insights for ESS companies through automatically deciphering patterns from complex data sources by using mathematical algorithms. Most commonly, ML models are trained using labelled data to learn the relationship between inputs and outputs for future predictions. ML is effective at generalization, offering a scalable alternative to physics-based models across different battery chemistries, use cases, and health states. ML tools can enhance the feasibility of safe, efficient, and economical operation of ESSs that can be applied across a fleet of EVs or for use under different load conditions.

The ML workflow involves data collection, pre-processing, model training, evaluation, and deployment. These are followed by continuous monitoring and refinement based on performance. Defining your modelling goals, acquiring appropriate data, and effectively conveying your results are the greatest challenges, but they are also the keys to opportunity in this field. However, there is a disconnect in the objectives of ML advancement in energy storage between academia and industry. In Figure 1, we break down the ML process and highlight the areas of focus between academia and industry. Academia tends to focus on innovative model development, adding novelty, and surpassing accuracy benchmarks. Industry targets return on investment, efficiency, and the practical advancements of the ML workflow. In our research with Li-ion battery modelling, we have studied the four major steps of ML and identified challenging areas and potential solutions from an academic perspective, while focusing on industry applications.

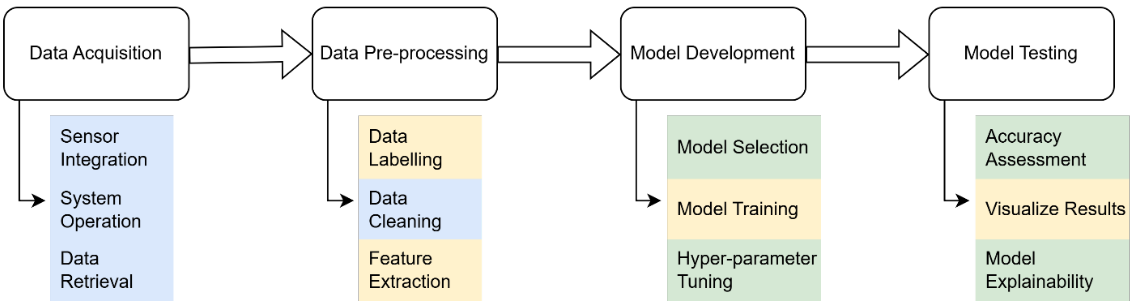

Figure 1: Major steps and sub-processes of machine learning coloured by focus of academia (green), industry (blue), and both (yellow)

Opportunities and barriers: Data is both the benefit and the challenge

Advances in computational resources, cloud storage, and communication hardware, paired with battery sensors, are facilitating the implementation of ML in energy storage. Battery operators apply this data to predict usable system capacity, monitor state of charge, forecast remaining useful life, and screen the performance of new battery chemistries (Li et al., 2022; Mao et al., 2021; Zhai et al., 2024). One of the most promising opportunities is forecasting electricity pricing to advise smart battery charging and discharging for energy arbitrage in dynamic electricity markets. Sensors integrated in ESSs are ideally suited to allow for the acquisition of large datasets usable for ML model development. However, proprietary data restrictions and laboratory hardware limitations make this uncommon from an academic standpoint. Collecting battery cycling data in the lab can take between 2-24 months, and with hardware/material limitations data quantity and quality are affected. As sufficient data is crucial to ML success, we have explored techniques to make the most of limited data.

In academic settings, feature extraction and data augmentation are common pre-processing methods to transform data in order to improve model performance without access to a large dataset. In contrast, industry focuses more on practical data pre-processing needs for large datasets, such as cleaning messy data and structuring long time-series from sensors. More recently the field of ML has been trending towards the use of pre-trained and foundational models such as ChatGPT and TabPFN which seamlessly synthesize new data based on a vast and diverse training set, rather than building bespoke models from scratch for each new dataset. With unified goals, between academia and industry, advancing these pre-trained models for ESS-specific cases could have significant benefits in fulfilling broadly shared requirements similar to the advent of chatbots built on a pre-trained large language model like ChatGPT.

Another challenge in applying ML is the human interpretation of the model's decision-making. The black-box nature of most ML tools does not foster trust between model users and the predictions made. Due to laboratory constraints, robust testing of academic methods under real-world conditions is limited. To make up for this, academia leans on transparency in reporting prediction accuracy metrics, confidence/uncertainty, and by using explainable ML techniques that directly expose how final outputs are arrived at by the ML algorithm and inputs.

Next steps: Merging machine learning goals with academia and industry

In summary there are two key challenges in applying machine learning to energy storage systems are: 1) the difficulty of collecting high-quality, long-term system data in laboratory settings; 2) the lack of transparency in how models make predictions. Bridging this gap between academia and industry through collaboration has the potential to bring advanced methods to real-world large dynamic datasets, benefiting both parties. Canada's strategic approach to battery innovation involves encouraging and enabling connections between industry, federal labs, and academics (Canada, 2024). With targeted policies and with incentives to support shared projects and data infrastructure, Canada is well-positioned to be a leader in data-driven, intelligent energy storage management.

References

Canada, N. R. (2024, October 8). *Strategic Approach to Battery Innovation*. https://natural-resources.canada.ca/science-data/science-research/strategic-approach-battery-innovation

Li, A. G., West, A. C., & Preindl, M. (2022). Towards unified machine learning characterization of lithium-ion battery degradation across multiple levels: A critical review. *Applied Energy*, *316*, 119030. https://doi.org/10.1016/j.apenergy.2022.119030

Mao, J., Miao, J., Lu, Y., & Tong, Z. (2021). Machine learning of materials design and state prediction for lithium-ion batteries. *Chinese Journal of Chemical Engineering*, *37*, 1–11. https://doi.org/10.1016/j.cjche.2021.04.009

Zhai, Q., Jiang, H., Long, N., Kang, Q., Meng, X., Zhou, M., Yan, L., & Ma, T. (2024). Machine learning for full lifecycle management of lithium-ion batteries. *Renewable and Sustainable Energy Reviews*, *202*, 114647. https://doi.org/10.1016/j.rser.2024.114647