Journal of The Electrochemical Society, 2021 **168** 030525

Perspective

Perspective—Combining Physics and Machine Learning to Predict Battery Lifetime

Muratahan Aykol,[1] Chirranjeevi Balaji Gopal,[1] Abraham Anapolsky,[1] Patrick K. Herring,[1] Bruis van Vlijmen,[2,3] Marc D. Berliner,[4] Martin Z. Bazant,[4,*] Richard D. Braatz,[4,*] William C. Chueh,[2,3,*] and Brian D. Storey[1,z]

[1]*Toyota Research Institute, Los Altos, California 94022, United States of America*
[2]*SLAC National Accelerator Laboratory, Menlo Park, California 94025, United States of America*
[3]*Stanford University, Stanford, California 94305, United States of America*
[4]*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States of America*

Forecasting the health of a battery is a modeling effort that is critical to driving improvements in and adoption of electric vehicles. Purely physics-based models and purely data-driven models have advantages and limitations of their own. Considering the nature of battery data and end-user applications, we outline several architectures for integrating physics-based and machine learning models that can improve our ability to forecast battery lifetime. We discuss the ease of implementation, advantages, limitations, and viability of each architecture, given the state of the art in the battery and machine learning fields.
 [DOI: 10.1149/1945-7111/abec55]

Manuscript submitted November 1, 2020; revised manuscript received January 25, 2021. Published March 16, 2021. *This paper is part of the JES Focus Issue on Selected Papers of Invited Speakers to IMLB 2020.*

For battery electric vehicles to exceed parity with internal combustion engines, substantial improvements in battery performance, cost, and longevity are needed. To make these improvements, one critical need is the capability to accurately forecast lithium-ion battery (LIB) lifetime and degradation.[1–5] Not only is forecasting important for guaranteeing the battery pack will last over 10 years, but it is also needed for accelerating the design cycle. Batteries are optimized over a complex space of materials (electrodes, electrolytes, binders, separators, etc.) and cell design parameters. This large design space coupled with wide variation in use conditions makes LIB system optimization challenging and slow. Since the lifetime evaluation process is a major bottleneck in design iterations, accurate prediction of the future state of health (e.g. remaining useful life; RUL) is needed to reduce the number of cells tested and length of each test.[6] Accurate prediction of RUL can also determine if a used cell should be re-purposed for a second-life application.[1]

Forecasting the future health of a battery is a multifaceted challenge.[1] Degradation results from an interplay of mechanisms that lead to loss of active Li-inventory or active electrode materials; decomposition of the solid electrolyte interphase (SEI), electrolyte, or binder; exfoliation of graphite anode; Li-plating/dendrite formation; cracking and loss of contact; and leaching and deposition of transition metals, gas formation, and other corrosive processes.[7] The health of a battery is not directly measurable and typically requires specific diagnostic cycles. In the field, we need to be able to project future health from partial charge/discharge cycles and the corresponding signals tracked by the battery management system (BMS) such as current, voltage, and temperature. Forecasting models must also generalize to future load and environmental conditions.

The ideal health forecasting approach would be based on *first-principles* models that can account for electrochemical and degradation processes. Physics-based (PB) models can compute the evolution of the internal states of the battery under an expected load and environment. The pseudo-two-dimensional (P2D) model, originally developed by Newman and co-workers,[8] is the most common framework for describing physical processes governing LIB systems. Due to the large number of parameters and computational complexity of the P2D model, researchers also turn to single-particle models or other simplified PB models to achieve a compromise between physical accuracy and complexity.[9–11] While PB models can capture the electrochemical cycling behavior, predicting long-term degradation is particularly challenging. Certain capacity degradation mechanisms such as SEI growth[12–15] and Li plating have been incorporated into PB models,[16–18] but many degradation modes remain poorly understood and existing degradation models suffer from poor identifiability of physical mechanisms.[19] Furthermore, the assumptions underlying common electrochemical transport equations break down at large driving forces, and different approaches are required based on nonequilibrium thermodynamics[20] and quantum theories of coupled ion-electron transfer.[21] There have been efforts to address some of these limitations of the Newman model, notably through the development of multiphase porous electrode theory (MPET),[22] but even the most advanced PB models still have limited applicability for health forecasting.

Recently, data-driven, machine learning-based (ML) models have shown success in predicting the RUL of Li-ion cells under various load conditions.[3,23–27] ML models operate by recognizing high-dimensional patterns in data, and are agnostic to the underlying physical processes. ML models can have orders of magnitude more free parameters than PB models and hence require large training datasets. They are also prone to unphysical failure when generalizing to scenarios unseen in training. To be able to learn from the available data, ML may also require *dimensionality reduction and feature engineering*, the latter of which is problem-specific.[26,28] ML is unlikely to bring about high-accuracy health forecasting transferable to situations far beyond the available data, without any consideration of physical processes.

There has been growing interest in blending physical and machine learning models to leverage their respective strengths in many fields including weather forecasting, biological systems, materials chemistry, mechanical failure, battery health, and battery safety.[2,4,5,29–31] The central question of this perspective is: how can integration of physics based models and machine learning models improve our ability to forecast battery lifetime? Fundamentally, degradation of LIBs results from evolution described by physical laws of thermodynamics and kinetics. However, real LIBs are complex (comprising multiple interfaces, materials and broad usage conditions) and successful models of battery degradation must inherently be capable of bridging spatial, temporal, and chemical complexity. Given the success of PB models in describing cell behavior and early-stage capacity fade, and the emergence of ML

*Electrochemical Society Member.
[z]E-mail: brian.storey@tri.global

models that generate rapid predictions of future health based on descriptors learned purely from data, novel approaches that hybridize the two could tap into both their advantages. In this perspective, we discuss several architectures to combine PB and ML models for battery health predictions, along with their relative ease of implementation and value in both the short and long term.

Current Status

There are three primary time-dependent data streams that describe a battery in operation: current (i), voltage (V), and temperature (T). The role of any modeling effort for predicting the future health of a battery is to map the time-evolution of these variables to a target performance metric (e.g. RUL, capacity at a certain future time) as the battery ages. While PB models operate naturally on these data, ML typically involves a "feature engineering" step to derive new features that correlate more strongly with the target (e.g. the $\Delta Q(V)$ features by Severson et al.[26]) We refer to such time-dependent input features collectively as the vector $\mathbf{X}(t)$ and the time dependent target health metrics as $\mathbf{y}(t)$. As common in time-series forecasting, if *on-the-fly* measurement of $\mathbf{y}$ is possible, the past values of $\mathbf{y}(t < t_{\text{present}})$ can also become inputs.

When batteries are cycled at specified rates in a laboratory, such health metrics can be directly assessed. For a vehicle in operation, there are no standard cycles and such metrics are not directly available. Since the target $\mathbf{y}$ has to be known in training, diagnostic cycles should be embedded in cycling experiments that mimic the real-world use or in the BMS procedures of the vehicle to extract the health metrics $\mathbf{y}$.[23] New methods which eliminate the need for a diagnostic cycle to determine a battery's health metric are also promising. Recent research has shown that an *adaptive observer* based on an electrochemical model can closely monitor the capacity over the course of a cell's degradation, by updating the model parameters as the cell degrades.[32] Such an approach could enhance the efficacy of datasets lacking diagnostic cycles, which would otherwise not be suitable for battery health prediction.

Several possible integration architectures for physics based and machine learning models are outlined in Fig. 1. At a high level, there are two broad categories for health forecasting: (A) serial integration of independent models and (B) hybridized PB and ML models. The former category involves architectures more viable in the near term as they can be realized by integration of existing ML and PB tools without any fundamental changes. The latter category will require the development of new approaches.

In architecture A1, the ML model learns the *residual* between the PB prediction and the experimental target. Baseline physics can be captured with a PB model while the more complicated degradation is learned by a flexible ML model. The PB models can range from fast, phenomenological models using physical scaling laws,[33] to the slower but more mechanistic single-particle model or models from the P2D framework.[9–11,34–37] The parameters of the PB model are fixed and not *learned* during training. Since the PB model precedes ML in training and prediction, the PB model needs to be lightweight in order to not introduce a bottleneck, limiting the practical choices to simpler PB reduced-order models. Reduced-order models have an advantage of being more physically identifiable and less susceptible to overfitting, but these simple models may also lack sufficient explanatory power for complex degradation prediction. Compared to pure ML, this architecture has its output partially constrained by physical laws so may better generalize to new battery designs. Since the parameters of the PB model are not learned from the data, this architecture may still have limitations in capturing the natural variety of experimental data.

Architecture A2 overcomes two problems: (i) low complexity requirement for the PB model by taking it out of forward prediction and (ii) exclusive reliance on experimental cycling data by expansion of training data using PB model. In A2 the PB model generates data analogous to experiments, preferably factoring in multiple degradation modes. The parameters of the PB model which generates synthetic data must have different sets of parameters to represent distinct experiments. The selected ML model is then trained on experimental and computational data, potentially allowing a more accurate ML model for the same experimental dataset. The A2 integration is a specific implementation of *transfer learning*,[38] using the PB model to generate supplemental knowledge[33] to help train the ML model on *small* experimental datasets. Success of this transfer learning approach hinges on the PB model capturing relevant, and sufficient physics. Examples of transfer learning have been reported for battery state and health estimation,[39–41] but the specific integration outlined in A2 has not been reported and remains a viable opportunity. As in A1, the output of A2 is produced by ML and hence is constrained only partially by the known physics.

In architecture A3, the concept is to learn *input parameters* of a PB model. A typical implementation may start with a PB reduced order model that has enough degradation physics to broadly capture the battery dynamics. Using experimental data as a target, the PB model is then *refitted* using a selected set of input parameters as variables ($\boldsymbol{\theta}_{\text{PBM}} = [\theta_1, \theta_2, \ldots, \theta_n]$ where θ denotes parameters such as diffusion coefficients, rate constants, and exchange-current density). In this framework, ML predicts $\boldsymbol{\theta}_{\text{PBM}}$ and the PB model uses $\boldsymbol{\theta}_{\text{PBM}}$ to predict $\mathbf{y}$. The primary advantages of A3 are that the prediction is fully constrained by the PB model and the features yield an interpretable framework. However, models such as P2D have a large number of parameters which makes proper parameter identification challenging and it may not be possible to prevent the model from learning unphysical parameters. If the requirement of the PB model is to only obtain a match to the experimental data then the physical identification of model parameters may not be important. While this sequential-integration approach may be difficult to train, A3 is likely to become one of the best performing integrated models to deploy. Examples of A3 architecture can be seen in other fields related to prognostics or forecasting.[29,42] An example for predicting LIB health was provided by Ramadesigan et al.[43] who proposed forecasting the capacity fade by extrapolation of effective diffusion and rate constants. Another example was showcased by Bills et al.,[44] where the changes in internal state variables of an electrochemical model due to aging are predicted by a neural network.

The sequential integration methods keep ML and PB modeling as distinct elements of the overall pipeline, and hence are implementable with existing methodologies. While we present three possible architectures, there are numerous other architectures that might be hybrids or variations on the options we put forth here. An alternative path for integration is hybridization of ML and PB modeling paradigms where the boundaries around each modeling paradigm become diffuse, either as physics-constrained ML (B1) or machine learning accelerated PB modeling (B2).

The architecture B1 is typically accomplished through either a physics-guided loss function and/or model architecture. Physics-based penalty terms reflecting energy conservation and density variations have been used to improve ML predictions and capture physical spatio-temporal relations in complex natural systems.[45,46] Borrowing from information-theoretic approaches, physics-based regularizers have been used to balance complexity and predictability of deep neural network models to forecast rare event dynamics of biomolecules.[47] The other emerging methodology for B1 is physics-guided design of network architecture, examples of which include crystal graph convolutional networks to represent materials, machine learning interatomic potentials and information-bottleneck approaches to learn low-dimensional, physical representations from deep neural networks.[30,47–50] A possible extension of B1 to battery health forecasting could be a generative ML model trained to learn the state of the battery, and the use of conservation laws to constrain the time-evolution of model weights/coefficients, an approach demonstrated to date on simple physical systems.[51,52] An alternative is to solve PDE-constrained inverse problems for the key constitutive relations in PB models from datasets, combining electrochemical cycling, spectroscopy, and microscopy.[53]

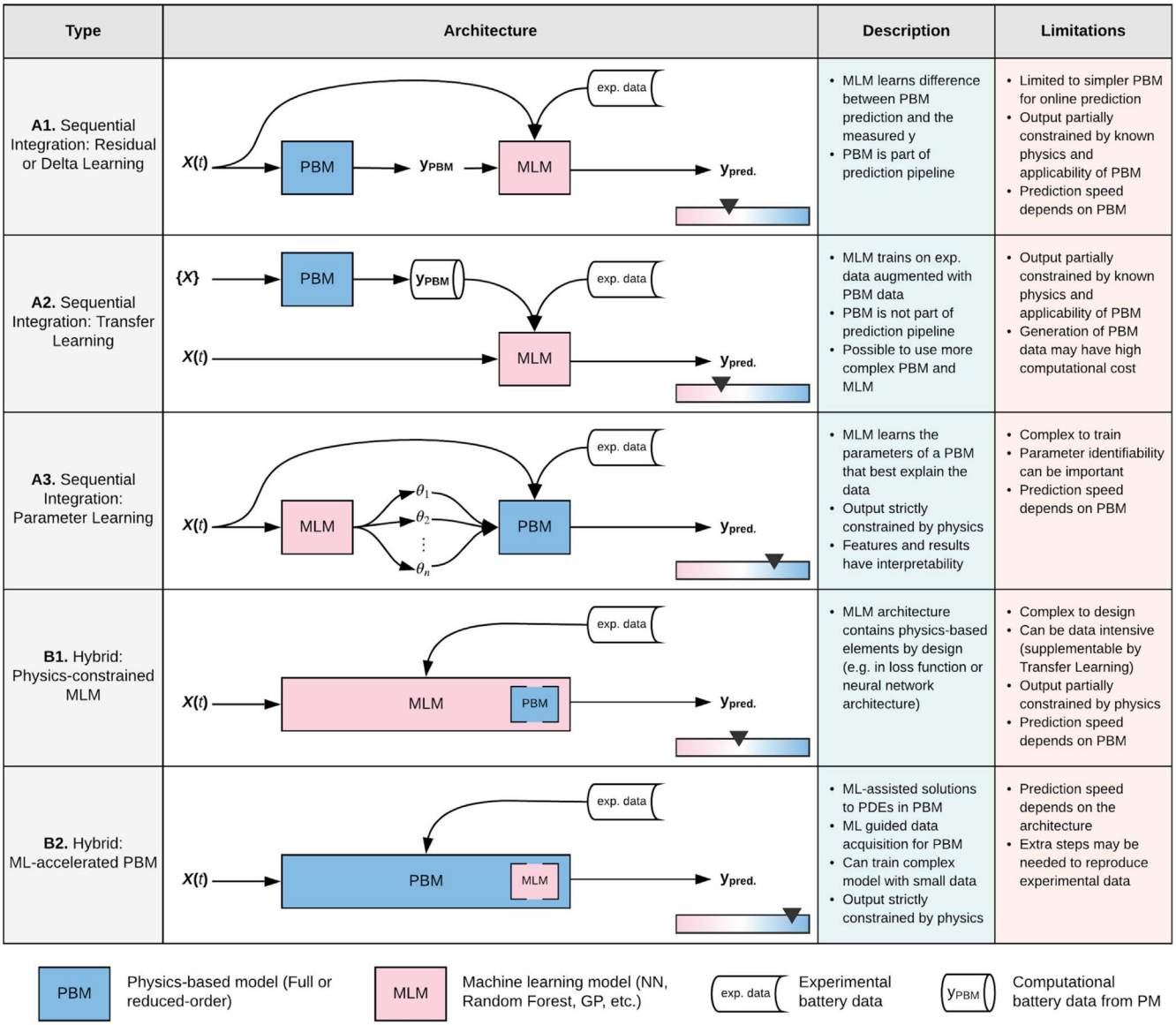

Figure 1. A map of integration strategies for physics-based and machine learning models for forecasting battery health. There are two broad forecasting categories; namely, A: Sequential Integration and B: Hybrid models. The architecture diagrams are drawn to depict primarily the prediction workflows. The entry points of experimental and computational training data are explicitly shown. Strategies for training each type of integration are described in the text. The color scales convey, relatively, the expected ratio of information captured by machine learning part vs physical part of an integrated model. In all schematics, time dependent input features are contained in the vector $\mathbf{X}(t)$ and the time-dependent target health metrics are $\mathbf{y}(t)$. $\boldsymbol{X}(t)$ could be any combination of raw time-series data (e.g. voltage, current, temperature), and/or any additional derived features. In A2, {X} denotes input parameter sets for the PB model defining use conditions (e.g. time vs current) analogous to that of experimental data to produce a dataset of simulated y_{PBM} values useful for transfer learning.

The architecture of type B2 relies on physics-informed ML designed to rapidly solve the underlying nonlinear PDEs for dynamic battery models, and potentially provide new physical insight under limited experimental measurements. Such hidden physics models have been pioneered by Raissi and co-workers,[54–57] leading to the design of physics-informed neural networks (PINNs).[54] PINNs exploit the automatic differentiation and universal function approximator aspects of neural networks to train on, solve and/or discover the nonlinear PDEs of the observed system, with small amounts of data, in effect obeying the underlying physical laws and boundary conditions.[54] This area is rapidly evolving, and numerous diverse examples in multiphysics modeling applications have been shown, such as in fluid mechanics, Lagrangian mechanics, phase-field modeling of fracture, and Schrodinger's equation.[54,57–59] The principal governing equations of battery modeling stem from species and charge conservation and electrochemical reactions, and become increasingly complex with inclusion of degradation mechanisms that activate at different stages of life and operating conditions. Extension of PINNs to electrochemical modeling of batteries is therefore a challenging but exciting frontier that can ultimately enable learning the hidden physics from experimental cycling and result in a significant advance in the field.

Future needs and prospects.—We expect all the strategies laid out in Fig. 1 for blending physics and machine learning to mature and find practical use in battery forecasting and other fields. Sequential integration of standalone PB and ML modeling tools (type A architectures) is a well-defined challenge focused primarily on software development and integration. Sequential architectures are solvable *today*, through integration, improvement, and repurposing of existing scientific software. On the other hand, the hybrid architectures (type B) for electrochemical modeling remains an open research question. Type A is a practical near-term strategy for

battery health forecasting, whereas type B will become dominant in the long term. The class of hybrid approaches have potential to fuse the causative and extrapolative capabilities of physics-based models with the speed, flexibility, and higher-dimensional capability of data-driven models, such as deep neural networks.

There are multiple challenges that need to be addressed to reduce the integrated PB-ML modeling architectures to practice for battery health forecasting. First, large systematic cycling datasets relevant for driving scenarios are needed to properly train, validate and benchmark all integrated models and improve their generalizability.[3,23,27] Research groups have started releasing datasets and related software libraries to address a shortage of publicly available battery data.[26,60–64] Publicly hosted datasets, together with well-maintained open-source packages[63,64] for standardizing battery data processing workflows can help accelerate all modeling efforts including those aimed toward better PB-ML integration. Quantitative benchmarks are necessary not just to determine improvements in predictions, but also evaluate speed and implementation cost of real-time prognostics. Second, producing large amounts of data under a variety of conditions from a PB model requires reliable computational workflows that can be run on-demand. Several open-source electrochemical modeling libraries are progressing towards this goal.[9,22,65,66] Finally, a concerted effort is needed for building integrated PB-ML models as a maintainable platform,[67] which can leverage the growth in infrastructure developed for streamlining ML workflows.

While battery cycling data are relatively straightforward to obtain for a large number of cells over a long period of operation, the data are limited to voltage, current, and temperature. Using characterization data that go beyond cycling can provide richer information for machine learning,[68] but is limited in the number of cells and conditions that can be tested. Multi-fidelity machine learning or other techniques will be needed to hybridize diverse data sets and ultimately improve lifetime forecasting. Advanced imaging techniques have been used to characterize a variety of mesoscale phenomena, and these data have dramatically improved our understanding and development of physics-based models.[53,69] High-fidelity data which allows inference of the details of the thermodynamic landscape of materials or reaction rates can improve the accuracy of current PB models and are also needed for development of hybrid models. Better performing PB models whose parameters are set by independent measurements should reduce the amount of fitting ML must ultimately supply.

What other battery problems can be tackled by integration of PB and ML models? One important class is integrating *online* control of cycling experiments to solve complex optimization problems. Near-term cost reductions in existing LIB technologies are expected to stem from optimizing design, manufacturing, and operation.[70,71] We expect ML-accelerated optimization to play an important role in driving these improvements just as such optimization is a promising frontier to accelerate materials design.[72–74] Optimization of a LIB over materials, additives, or system parameters can be facilitated in a closed-loop process where ML plays the role of a surrogate model which chooses the optimal set of design or use variables.[3,6,27,75] Attia et al.[6] demonstrated ML-driven closed-loop optimization of fast-charging protocols, where a forecasting model was used to reduce the required experimental cycling time by a factor of 30.[26] Optimization could also use a multi-fidelity setting,[76] where ML chooses from performance measurements of two fidelities; namely, PB model vs cycling experiment, to reduce cost.

For these complex optimization problems, we expect a few near-term opportunities. First, one opportunity is maintaining the current levels of ML prediction accuracy (>90%[6]) but integrating PB models to enable more complex design targets than have been achieved to date. One example is optimizing formation cycling which involves adjusting numerous protocol parameters. Through integration of physics, we expect improvements in model performance under unseen parameters which could reduce the overall number of parameters tested to optimize for battery lifetime while reducing manufacturing cost. We see a factor of 5 reduction in parameter searches from traditional approaches as a viable near-term target. Second, we expect that through integrating PB and ML models in order to improve overall prediction accuracy we can reduce the number of experimental cycles required to predict lifetime. Maintaining accuracy of current approaches but using only the first 10 cycles of experimental data is a viable near-term target. Finally, in many cases we need categorical screening such as finding a group of promising charging policies or sorting as-manufactured batteries. For categorization, current ML-based models tend to display high performance after only 5 cycles of data.[26] Successful integration of PB modeling could gain further speed-up which would have immediate use during manufacturing. All targets above are estimated using the acceleration demonstrated by Attia et al.[6] and Severson et al.[26] as baselines. For all these applications, we envision a positive feedback loop between improved forecasting, reduced optimization time, and more accurate physical models, ultimately leading to substantial acceleration of battery development.

Finally, the benefits of improved on-board lifetime prediction for a BMS are broadly discussed in the literature.[1,2,23,24,28] The use of integrated PB-ML models for on-vehicle applications is another promising near-term application, though the requirements for model accuracy are different than in design optimization, and variations in cycling conditions, forecasting horizons and computational costs make a cross-comparison of current model accuracies challenging. Nevertheless, if lifetime prediction methods are being used in the BMS to predict battery safety or reliability,[77] then the cost of missing a failure can be very high (the financial cost, safety cost, and the cost to the brand) whereas misclassifying some good batteries as bad is tolerable to a degree.

Conclusions

There has been notable progress in projecting the behavior of rechargeable batteries through two distinct modeling paradigms: ML and PB electrochemical modeling. ML offers a flexible *statistical* framework that yields rapid predictions but requires problem-specific feature design and lacks physical description of degradation modes, restricting its reliability to conditions covered by the training data. Physics-based models provide a *mechanistic* framework to capture the influence of degradation physics on electrochemical behavior, but challenges include poorly understood and hard-to-identify degradation modes. It has become clear in recent years that ML and PB models offer qualities distinct but complementary to each other, and a natural next phase is exploring their proper blending for achieving better battery lifetime prediction. We outlined several key strategies for integrating physics and ML via sequential architectures (type A) and hybrid architectures (type B), and discussed the advantages, limitations as well as short- and long-term feasibility of each architecture. Sequential architectures are implementable today using existing ML and PB tools as building blocks, but the current lack of available experimental data may still limit their performance and impact. Hybrid architectures still require extensive development to create physics-informed ML or ML-accelerated PB models. The promise of these hybrid architectures to extrapolate outside the training data means their development will likely accelerate. All architectures have the potential to overcome limitations of ML and PB approaches and improve upon the accuracies achievable through either modeling paradigm alone for battery health forecasting.

ORCID

Brian D. Storey https://orcid.org/0000-0001-5612-3390

References

1. X. Hu, L. Xu, X. Lin, and M. Pecht, *Joule*, **4**, 310 (2020).
2. L. Liao and F. Köttig, *IEEE Trans. Reliab.*, **63**, 191 (2014).

3. M.-F. Ng, J. Zhao, Q. Yan, G. J. Conduit, and Z. W. Seh, *Nature Machine Intelligence*, **2**, 161 (2020).
4. D. An, N. H. Kim, and J.-H. Choi, *Reliab. Eng. Syst. Saf.*, **133**, 223 (2015).
5. H. Meng and Y.-F. Li, *Renewable Sustainable Energy Rev.*, **116**, 109405 (2019).
6. P. M. Attia et al., *Nature*, **578**, 397 (2020).
7. C. R. Birkl, M. R. Roberts, E. McTurk, P. G. Bruce, and D. A. Howey, *J. Power Sources*, **341**, 373 (2017).
8. M. Doyle, T. F. Fuller, and J. Newman, *J. Electrochem. Soc.*, **140**, 1526 (1993).
9. J. M. Reniers, G. Mulder, and D. A. Howey, *J. Electrochem. Soc.*, **166**, A3189 (2019).
10. A. Jokar, B. Rajabloo, M. Désilets, and M. Lacroix, *J. Power Sources*, **327**, 44 (2016).
11. P. W. C. Northrop et al., *J. Electrochem. Soc.*, **161**, E3149 (2014).
12. B. Horstmann, F. Single, and A. Latz, *Current Opinion in Electrochemistry*, **13**, 61 (2019).
13. F. Single, B. Horstmann, and A. Latz, *J. Electrochem. Soc.*, **164**, E3132 (2017).
14. S. Das, P. M. Attia, W. C. Chueh, and M. Z. Bazant, *J. Electrochem. Soc.*, **166**, E107 (2019).
15. M. B. Pinson and M. Z. Bazant, *J. Electrochem. Soc.*, **160**, A243 (2012).
16. X.-G. Yang, Y. Leng, G. Zhang, S. Ge, and C.-Y. Wang, *J. Power Sources*, **360**, 28 (2017).
17. T. Waldmann, B.-I. Hogg, and M. Wohlfahrt-Mehrens, *J. Power Sources*, **384**, 107 (2018).
18. D. R. Baker and M. W. Verbrugge, *J. Electrochem. Soc.*, **167**, 013504 (2019).
19. U. Krewer, F. Roder, E. Harinath, R. D. Braatz, B. Bedurftig, and R. Findeisen, *J. Electrochem. Soc.*, **165**, A3656 (2018).
20. M. Z. Bazant, *Acc. Chem. Res.*, **46**, 1144 (2013).
21. P. Bai and M. Z. Bazant, *Nat. Commun.*, **5**, 3585 (2014).
22. R. B. Smith and M. Z. Bazant, *J. Electrochem. Soc.*, **164**, E3291 (2017).
23. G.-W. You, S. Park, and D. Oh, *Appl. Energy*, **176**, 92 (2016).
24. A. Nuhic, T. Terzimehic, T. Soczka-Guth, M. Buchholz, and K. Dietmayer, *J. Power Sources*, **239**, 680 (2013).
25. Y. Li, K. Liu, A. M. Foley, A. Zulke, M. Berecibar, E. Nanini-Maury, J. Van Mierlo, and H. E. Hoster, *Renewable Sustainable Energy Rev.*, **113**, 109254 (2019).
26. K. A. Severson et al., *Nat. Energy*, **4**, 383 (2019).
27. M. Aykol, P. Herring, and A. Anapolsky, *Nature Reviews Materials*, **5**, 725 (2020).
28. R. R. Richardson, M. A. Osborne, and D. A. Howey, *Journal of Energy Storage*, **23**, 320 (2019).
29. J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar, *arXiv [physics.comp-ph]* (2020), (http://arxiv.org/abs/2003.04919).
30. T. Xie and J. C. Grossman, *Phys. Rev. Lett.*, **120**, 145301 (2018).
31. G. C. Y. Peng et al., *Arch. Comput. Methods Eng.* (2020).
32. A. Allam and S. Onori, *IEEE Trans. Control Syst. Technol.*, 1 (2020).
33. M. Dubarry and D. Beck, *Mathematics & Computer Science* (2020), (https://preprints.org/manuscript/202005.0106/v1).
34. V. R. Subramanian, V. Boovaragavan, and V. D. Diwakar, *Electrochem. Solid-State Lett.*, **10**, A255 (2007).
35. S. K. Rechkemmer, X. Zang, W. Zhang, and O. Sawodny, *Journal of Energy Storage*, **21**, 773 (2019).
36. W. He, M. Pecht, D. Flynn, and F. Dinmohammadi, *Energies*, **11**, 2120 (2018).
37. S. J. Moura, F. B. Argomedo, R. Klein, A. Mirtabatabaei, and M. Krstic, *IEEE Trans. Control Syst. Technol.*, **25**, 453 (2017).
38. K. Weiss, T. M. Khoshgoftaar, and D. Wang, *Journal of Big Data*, **3**, 9 (2016).
39. S. Shen, M. Sadoughi, and C. Hu, in*2019 IEEE Transportation Electrification Conference and Expo (ITEC)*, p. 1 (2019).
40. S. Shen, M. Sadoughi, M. Li, Z. Wang, and C. Hu, *Appl. Energy*, **260**, 114296 (2020).
41. C. Vidal, P. Kollmeyer, E. Chemali, and A. Emadi, in*2019 IEEE Transportation Electrification Conference and Expo (ITEC)*, p. 1 (2019).
42. M. A. Chao, C. Kulkarni, K. Goebel, and O. Fink, *arXiv [eess.SY]* (2020), (http://arxiv.org/abs/2003.00732).
43. V. Ramadesigan et al., *J. Electrochem. Soc.*, **158**, A1048 (2011).
44. A. Bills, S. Sripad, W. L. Fredericks, M. Guttenberg, D. Charles, E. Frank, and V. Viswanathan, *arXiv [physics.app-ph]* (2020), (http://arxiv.org/abs/2008.01527).
45. X. Jia, J. Willard, A. Karpatne, J. S. Read, J. A. Zwart, M. Steinbach, and V. Kumar, *arXiv [cs.LG]* (2020), (http://arxiv.org/abs/2001.11086).
46. A. Karpatne et al., *IEEE Trans. Knowl. Data Eng.*, **29**, 2318 (2017).
47. Y. Wang, J. M. L. Ribeiro, and P. Tiwary, *Nat. Commun.*, **10**, 3573 (2019).
48. R. Iten et al., *Phys. Rev. Lett.*, **124**, 010508 (2020).
49. V. L. Deringer, M. A. Caro, and G. Csányi, *Adv. Mater.*, **31**, e1902765 (2019).
50. N. Artrith and A. Urban, *Comput. Mater. Sci.*, **114**, 135 (2016).
51. P. Toth, D. J. Rezende, A. Jaegle, S. Racaniere, A. Botev, and I. Higgins, *arXiv [cs.LG]* (2019), (http://arxiv.org/abs/1909.13789).
52. S. Greydanus, M. Dzamba, and J. Yosinski, *Advances in Neural Information Processing Systems 32*, ed. H. Wallach et al. (Curran Associates, Inc.)15379 (2019), https://arxiv.org/abs/1906.01563.
53. H. Zhao, B. D. Storey, R. D. Braatz, and M. Z. Bazant, *Phys. Rev. Lett.*, **124**, 060201 (2020).
54. M. Raissi, P. Perdikaris, and G. E. Karniadakis, *arXiv [cs.AI]* (2017), (http://arxiv.org/abs/1711.10561).
55. M. Raissi and G. E. Karniadakis, *J. Comput. Phys.*, **357**, 125 (2018).
56. M. Raissi, *J. Mach. Learn. Res.*, **19**, 1 (2018).
57. M. Raissi, A. Yazdani, and G. E. Karniadakis, *Science*, **367**, 1026 (2020).
58. M. Lutter, C. Ritter, and J. Peters, *arXiv [cs.LG]* (2019), (http://arxiv.org/abs/1907.04490).
59. S. Goswami, C. Anitescu, S. Chakraborty, and T. Rabczuk, *arXiv [stat.ML]* (2019), (http://arxiv.org/abs/1907.02531).
60. Edward F. Hogge, Brian M. Bole, Sixto L. Vazquez, and Jose Celaya, *Verification of a Remaining Flying Time Prediction System for Small Electric Aircraft*, *Annual Conference of the Prognostics and Health Management, PHM 2015* (2015), https://www.phmsociety.org/node/1649.
61. S. Saxena, C. Hendricks, and M. Pecht, *J. Pow. Sour.*, **327**, 394-400 (2016).
62. Battery Archive, (http://batteryarchive.org/).
63. P. Herring et al., *SoftwareX*, **11**, 100506 (2020).
64. S. Buteau, (2020), Dahn Group at Dalhousie University https://github.com/Samuel-Buteau/universal-battery-database.
65. V. Sulzer, S. G. Marquis, R. Timms, M. Robinson, and S. J. Chapman, (2020), (https://ecsarxiv.org/67ckj/).
66. M. Torchio, L. Magni, R. Bhushan Gopaluni, R. D. Braatz, and D. M. Raimondo, *J. Electrochem. Soc.*, **163**, A1192 (2016).
67. M. Aykol et al., *Matter*, **1**, 1433 (2019).
68. Y. Zhang, Q. Tang, Y. Zhang, J. Wang, and U. Stimming, *Nat. Commun.*, **11**, 1706 (2020).
69. J. Lim et al., *Science*, **353**, 566 (2016).
70. R. Schmuch, R. Wagner, G. Hörpel, T. Placke, and M. Winter, *Nat. Energy*, **3**, 267 (2018).
71. B. Nykvist and M. Nilsson, *Nat. Clim. Chang.*, **5**, 329 (2015).
72. D. Xue, P. V. Balachandran, J. Hogden, J. Theiler, D. Que, and T. Lookman, *Nat. Commun.*, **7**, 11241 (2016).
73. J. H. Montoya, K. T. Winther, R. A. Flores, T. Bligaard, J. S. Hummelshoj, and M. Aykol, *Chem. Sci.*, **11**, 8517 (2020).
74. D. P. Tabor et al., *Nature Reviews Materials*, **3**, 5 (2018).
75. A. Dave, J. Mitchell, K. Kandasamy, S. Burke, B. Paria, B. Poczos, J. Whitacre, and V. Viswanathan, *arXiv [physics.app-ph]* (2019), (http://arxiv.org/abs/2001.09938).
76. K. Kandasamy, G. Dasarathy, J. B. Oliva, J. Schneider, and B. Póczos, *Advances in Neural Information Processing Systems*(Current Associates Inc, Red Hook, New York, United States of America)**29**, 992 (2016), https://arxiv.org/abs/1603.06288.
77. D. P. Finegan, J. Zhu, X. Feng, M. Keyser, M. Ulmefors, W. Li, M. Z. Bazant, and S. J. Cooper, *Joule*, **5**, 316 (2020).