Cost, Performance Prediction and Optimization on Vanadium Flow Battery by Machine-Learning

Article *in* Energy & Environmental Science · September 2020

DOI: 10.1039/D0EE02543G

7 authors, including:

Tianyu Li
Dalian Institute of Chemical Physics
16 PUBLICATIONS **22** CITATIONS
SEE PROFILE

Tao Liu
Dalian Institute of Chemical Physics
23 PUBLICATIONS **661** CITATIONS
SEE PROFILE

Xianfeng Li
Chinese Academy of Sciences
289 PUBLICATIONS **7,616** CITATIONS
SEE PROFILE

Energy & Environmental Science

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d0ee02543g

Cost, performance prediction and optimization of a vanadium flow battery by machine-learning†

Tianyu Li, Feng Xing, Tao Liu, Jiawei Sun, Dingqin Shi, Huamin Zhang and Xianfeng Li *

Received 9th August 2020,
Accepted 21st September 2020

DOI: 10.1039/d0ee02543g

rsc.li/ees

Performance optimization and cost reduction of a vanadium flow battery (VFB) system is essential for its commercialization and application in large-scale energy storage. However, developing a VFB stack from lab to industrial scale can take years of experiments due to the influence of complex factors, from key materials to the battery architecture. Herein, we have developed an innovative machine learning (ML) methodology to optimize and predict the efficiencies and costs of VFBs with extreme accuracy, based on our database of over 100 stacks with varying power rates. The results indicated that the cost of a VFB system (S-cost) at energy/power (E/P) = 4 h can reach around 223 $ (kW h)$^{-1}$, when the operating current density reaches 200 mA cm^{-2}, while the voltage efficiency (VE) and utilization ratio of the electrolyte (UE) are maintained above 90% and 80%, respectively. This work highlights the potential of the ML methodology to guide stack design and optimization of flow batteries to further accelerate their commercialization.

Broader context

The widespread application of renewable energies such as wind and solar power calls for economical and efficient energy-storage systems owing to their inherent intermittent nature. Flow batteries, especially vanadium flow batteries (VFBs), are among the most attractive technologies for large-scale energy storage, owing to their flexible design, high safety, high efficiency and long cycle life. Currently, VFBs are at a commercial demonstration stage; however, the relatively high cost restricts their further commercialization. Normally, the performance of a VFB stack varies with the materials, cell architecture and operation conditions, while the cost of a VFB system is highly dependent on its performance. It requires considerable effort and time to optimize the materials and structures of a VFB stack by experiment. Herein, we applied an innovative machine learning (ML) methodology to optimize and predict the performance and cost of a VFB system with extremely high accuracy, based on a database of over 100 stacks. Based on the ML model, the main factors that affect the battery performance and cost were clarified and the way to optimize the system architecture was provided. This work highlights the prospects of combining experimental data with statistics and ML algorithms to assist optimization and performance prediction of flow battery systems.

Introduction

The widespread application of renewable energies such as wind and solar power calls for economical and efficient energy-storage systems owing to their inherent intermittent nature.[1] In recent decades, various energy-storage technologies have been developed. These include physical methods such as pumped hydro and compressed air energy storage systems and electrochemical energy storage (EES) technologies.[2] Flow batteries, especially vanadium flow batteries (VFBs), are among the most attractive technologies for large-scale energy storage, owing to their flexible design, safety, efficiency, and long cycle life.[3,4] The performance and cost of a VFB system are highly related to the stack and electrolyte. Numerous efforts have explored high-performance key materials (membranes,[5–12] electrodes,[13–17] and bipolar plates[18]) and the optimization of the stack architecture including design of flow fields, optimizing the electrode compression ratio, and porosity.[19–22] Thus, the cost of VFB systems has lowered accordingly. Currently, VFBs are at a commercial demonstration stage; different projects ranging from kilowatts to over 100 megawatts have been implemented for different applications. However, the relatively high cost restricts further commercialization. Normally, the cost of a VFB system generally includes the power cost (P-cost) and energy cost (E-cost). The P-cost of a stack is determined by

Division of Energy Storage, Dalian National Laboratory for Clean Energy (DNL), Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Zhongshan Road 457, Dalian 116023, China. E-mail: lixianfeng@dicp.ac.cn

† Electronic supplementary information (ESI) available: Dataset-1, Dataset-2 and additional details including correlation analysis, residual plots of "only current density" models, accuracy evaluations of "only current density" models *etc.* See DOI: 10.1039/d0ee02543g

the stack performance. However, although the E-cost is closely related to the UE, it is also related to the stack performance, since the UE is correlated with the voltage efficiency of a stack. Normally, the performance of a VFB stack varies with the materials, architecture, and operation conditions, while the cost of a VFB system is highly dependent on its performance. It requires considerable effort and time to optimize the materials and structures of a VFB stack by experiment. Therefore, exploiting novel methods to accurately predict the performance and cost of a VFB stack and further systems is crucial to accelerate the research and development (R&D) of VFBs and their further commercialization.

Currently, the R&D of VFB stacks, including key materials, stack structure and performance optimization, is based on simulation, design, and experiments.[19–31] This approach is time consuming and expensive; moreover, it is difficult to clarify the influence of various factors on the system performance and cost. Previously, some techno-economic assessment models for the VFB system were established to assess the performance and cost. For example, Minke *et al.* performed a comprehensive techno-economic assessment of VFBs by combining the cost of key materials like electrodes[32] and membranes,[33] and proposed a transparent cost model for megawatt-size VFB systems[34] by dividing the system into P-cost and E-cost. Jens *et al.*[35] proposed an approach to calculate specific system costs and created a cost breakdown for a typical 10 kW/120 kW h vanadium redox battery (VRB) system. Li *et al.*[36] performed an economic analysis of the VRB system based on their improvement in electrolytes and cell stacks. The techno-economic assessment models have shown certain guiding significance for commercialization of VFBs; however, none of those models were based on experimental data with various operating conditions (*i.e.*, an operating current density range of 80 to 300 mA cm^{-2}) for VFB stacks. Therefore, it is vital for the R&D of VFB stacks and their rapid commercialization to develop an effective methodology that can accurately predict the stack performance and further dynamically evaluate the cost of a VFB system.

In the past decade, ML approaches have shown great potential in accelerating material design,[37–41] planning chemical synthesis,[42,43] tuning catalyst activity,[44–48] and optimizing systems,[49,50] and can thus predict and optimize performance in an efficient way. For example, Severson *et al.*[51] established a model that accurately predicted the cycle life of commercial lithium iron phosphate/graphite cells using early-cycle data through high-throughput experiments and ML. Later, Attia *et al.*[52] developed and demonstrated a closed-loop optimization of fast-charging protocols, which efficiently optimized a parameter space specifying the current and voltage profiles for maximizing the battery cycle life through ML. However, very rarely have studies focused on the prediction and optimization of VFB systems by ML, as it is difficult to collect sufficient parallel data on stacks with different parameters. In addition, most current research of flow batteries is focused on a single cell, which is atypical in real applications. To date, there is no model that can accurately predict the relationship between the materials and structures of VFB stacks, the operation conditions, the stack performance and the dynamic performance-cost of the system. This is due to the complexity of the influencing factors among the high-dimensional data and the discontinuity of basic R&D and industrialization research.

In the last two decades, our group has conducted fundamental and applied research on VFBs. Where many stacks were assembled, sufficient data is available for the ML methodology. Here, an ML framework that can assist in R&D of a high-performance VFB stack and accurately predict the relationship between the operating parameters and performance-cost of each VFB stack and system was built based on our R&D accumulation in VFB fields. The operating current density was used as the main feature. To characterize the VE, energy efficiency (EE), UE and E-cost, another 18 features were used, while a further 21 features were used to characterize the P-cost. Our ML models can achieve mean absolute prediction error (MPAE) within 1% for quantitatively predicting the VE and EE, and within 5.2% for the UE, P-cost and E-cost. Moreover, the P-cost, E-cost and S-cost of the optimized stack are lowered by 2.98%, 2.49% and 2.69%, respectively, compared to those of the original stack named ''VFB 20190225'' under the optimal operating conditions. Furthermore, the optimal operating S-cost in 2019 was reduced by 43.79% compared with that in 2013. Based on currently available technology, it is predicted that the S-cost of a VFB system can be reduced to $223 per kW h (at E/P = 4 h) under 200 mA cm^{-2} with VE = 90% and UE = 80%. This work highlights the prospects of combining experimental data with statistics and an ML algorithm to assist performance optimization and prediction of future flow battery systems.

Results and discussion

Methods

Performance test of VFB stacks. All materials, including their structure parameters and price, for each stack are listed in Dataset-1 (the ''latest'' dataset, which means that the data was collected from 2018 to 2020) and Dataset-2 (the ''earlier'' dataset, which means that the data was collected from 2012 to 2018). The exchange rate of 1 Dollar to Ren Min Bi (RMB) is settled at 7.1. The type 1 bipolar plate (BP Type 1) and membrane (M Type 1) were self-made. The test was performed under a constant charging and discharging current with an Arbin 2000. The charging cutoff voltage of a single cell is given in the column named ''Charging Cut-off Voltage (V/Cell)'' in each dataset, which was 1.55 V and 1.58 V, respectively. The P-cost, E-cost, and S-cost are the costs calculated from the raw materials for the VFB stacks and electrolytes in this study. The P-cost is determined by the stack and the E-cost is from the electrolytes. Note that the cost of pipes and the control system is not included.

Machine-learning models and evaluation criterion. The data processing and ML modeling are performed with Python 3.7 using the NumPy, pandas and sklearn packages. Linear regression was applied to handle the multidimensional data.

First, one-hot encoding was used to handle the category parameters such as the electrode, bipolar plate, and flow field, and then a regularization technique was applied to the numerical dataset. A linear regression model can be calculated by

$$f(\mathbf{X}, \mathbf{w}, b) = \mathbf{w}^T\mathbf{X} + b \quad (1)$$

where $\mathbf{X} = \langle x_1, x_2, \ldots, x_n\rangle$ is the n-dimensional feature vector. $\mathbf{w} = \langle w_1, w_2, \ldots, w_n\rangle$ is the corresponding coefficient vector of $\mathbf{X}$, and b is the intercept. Here, $f(\mathbf{X}, \mathbf{w}, b)$ can be VE (%), EE (%), UE (%), P-cost, and E-cost, respectively. The generalized least squares method was employed to minimize the loss function $L(\mathbf{w}, b)$, and the objective formula is as follows,

$$\arg\min_{\mathbf{w},b} L(\mathbf{w}, b) = \arg\min_{\mathbf{w},b} \sum_{i=1}^{m} \left(f_i(\mathbf{X}, \mathbf{w}, b) - y^i\right)^2 \quad (2)$$

where $L(\mathbf{w}, b)$ is the loss function, $\mathbf{y} = \langle y^1, y^2, \ldots, y^m\rangle$ is the target function, m is the number of data in the dataset, and y^i is one component in $\mathbf{y}$.

To evaluate the ML prediction performance, three standards are followed:

(1) Coefficient of determination (R^2)

$$R^2 = 1 - \frac{\mathrm{RSS}}{\mathrm{TSS}} = 1 - \frac{\sum_i (y^i - f_i)^2}{\sum_i (y^i - \hat{y})^2} \quad (3)$$

where RSS is the residual sum of squares and TSS is the total sum of squares. f_i is the prediction value of $\boldsymbol{x}_i$ and $\hat{y}$ is the average value of the target function.

(2) Root mean square error (RMSE)

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_i \left(f_i - \hat{f}\right)^2} \quad (4)$$

where $\hat{f}$ is the average value of the predicted results.

(3) Mean absolute prediction error (MAPE)

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{f_i - y^i}{y^i}\right| \times 100\% \quad (5)$$

ML is a general term, which is based on data-driven models and statistics. When we have enough effective raw data and give the computer an association rule discovery algorithm, the computer has the potential to discover unknown correlations between the input and output, which is also known as the "training" process. We can then use this trained model to predict the quantitative correspondence between the multiple input features and target output variables. Here an operating parameter-based (operating current density) method was applied to build the performance and cost prediction model. According to Fig. S1(a) (ESI†), there is a strong linear correlation between the operating current density and stack efficiencies, while the relationship between the operating current density and cost is near an exponential form [Fig. S1(b), ESI†]. After calculating the natural logarithm of P-cost and E-cost (they are referred to as adjusted costs in the following), there is also a linear correlation between the operating current density and adjusted cost. ln(power cost – 26) and ln(energy cost – 80) were chosen as the target function because the maximum absolute relative deviation (MARD) of ln(power cost – 26) and ln(energy cost – 80) is almost the lowest, and the MARD between the training set and test set is basically consistent (see Table S1, ESI†). Although linear regression is a traditional tool in mathematics and engineering, it has a unique advantage, that is, the interpretability between the features and the result is excellent. Therefore, an operating parameter-based ML algorithm was applied to predict the stack efficiencies and cost.

Accuracy evaluations of operating parameter-based ML models

In linear regression, the selection of appropriate features is especially important. Since the operating current density has a strong linear correlation with the stack efficiencies and adjusted costs, the operating current density is regarded as the main feature. All materials and some main structures which are considered to have a significant influence on the performance of the stack are then used as the other auxiliary features. All features for each model are shown in the datasets.

After building Dataset-1, we used 75% of the data as the training set, and the remaining 25% as the test set to evaluate the predictive accuracies of the model. The accuracy evaluations of our "full features" models for Dataset-1 are shown in Table 1, and the residual plots are shown in Fig. 1(a)–(g). The accuracy evaluations of our "only current density" models for Dataset-1 are shown in Table S2 (ESI†) and the residual plots are shown in Fig. S2(a)–(g) (ESI†). The comparison of the MAPE

Table 1 Accuracy evaluations of the "full features" models. The coefficient of determination (R^2), root mean square error (RMSE) and mean absolute prediction error (MAPE) are used to evaluate the precision of the linear regression models that are used to predict VE, EE, UE, ln(power cost – 26), ln(energy cost – 80), P-cost, and E-cost

	R^2		RMSE		MAPE (%)	
"Full features" models	Training	Test	Training	Test	Training	Test
VE (%)	0.9808	0.9848	0.5989	0.5141	0.54	0.47
EE (%)	0.9765	0.9779	0.6495	0.5920	0.62	0.56
UE (%)	0.9627	0.9494	2.2639	2.4031	3.40	3.14
ln(power cost – 26)	0.9715	0.9740	0.0840	0.0813	5.14(P-cost)	4.97(P-cost)
ln(energy cost – 80)	0.9727	0.9636	0.0659	0.0688	3.22(E-cost)	3.08(E-cost)
Power cost	0.8744	0.8859	23.3839	21.5280	12.47	11.22
Energy cost	0.9322	0.9220	16.7807	14.4318	5.69	7.31

The unit of P-cost and E-cost is ($ (kW h)$^{-1}$ at E/P = 4 h).

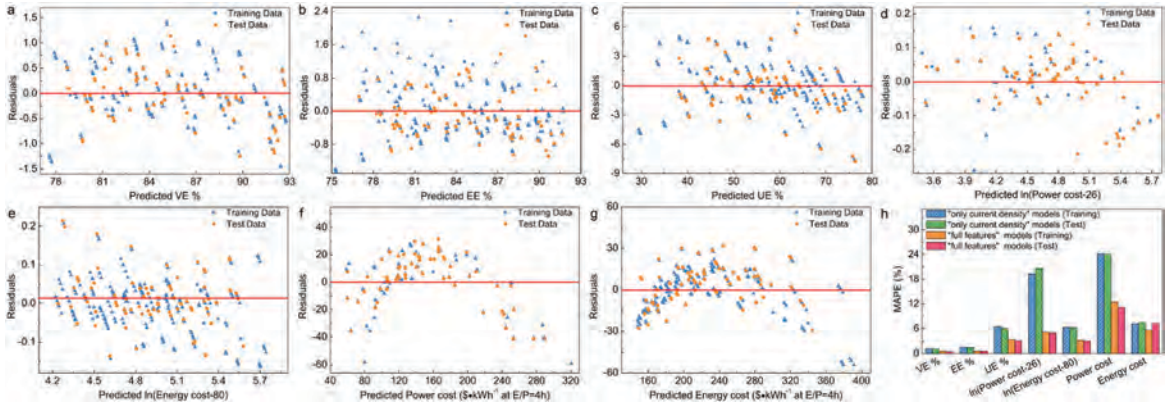

Fig. 1 The residual plots of the "full features" models for Dataset-1. (a) VE %, (b) EE %, (c) UE %, (d) ln(power cost − 26), (e) ln(energy cost − 80), (f) P-cost and (g) E-cost. The unit of P-cost and E-cost is ($ (kW h)$^{-1}$ at E/P = 4 h). (h) Comparison of the mean absolute prediction error (MAPE) between the "only current density" models and "full features" models.

between the "full features" models and the "only current density" models is shown in Fig. 1(h). The MAPE (training/test) of the "full features" models for VE, EE, UE, P-cost and E-cost are 0.54%/0.47%, 0.62%/0.56%, 3.40%/3.14%, 5.14%/4.97%, and 3.22%/3.08%, respectively, which are far more accurate than those of the "only current density" models. The results indicate the significance of the ML methodology and lay a good foundation for analysing the relationship between the performance and S-cost of VFB systems.

Comparison of results between experiment and "full features" models for Dataset-1

Based on our "full features" models, we predicted the efficiencies and cost variation with the operating current densities ranging from 100 mA cm^{-2} to 400 mA cm^{-2} of each VFB stack in Dataset-1. The results are shown in Fig. 2(a)–(h) and Fig. S3 (ESI†). For the stacks with a power range from 0.5 kW to 10 kW, the "full features" models have high prediction accuracy for both efficiencies and costs. Overall, most of the predicted values from the "full features" models are in agreement with the experimental ones. After comparing the predicted with the experiment results, it can be concluded that our "full features" models for Dataset-1 are reliable.

Optimization and design of the stack based on the "full features" models for Dataset-1

As mentioned above, the unique advantage of linear regression is that the interpretability between the features and the results is extremely high. After determining the validity of the "full features" models, we analyzed the coefficients for each model to determine the effect of each feature on the efficiencies or costs. The results are shown in Fig. 3(a) and (b). During data preprocessing, the category parameters were encoded by one-hot encoding, so the sum of coefficients of the same category is zero and the relative value of the coefficient can be used to determine its influence on the target function (VE, EE, UE, P-cost, and E-cost). From Fig. 3(a), it can be concluded that the effective way to lower the P-cost is to

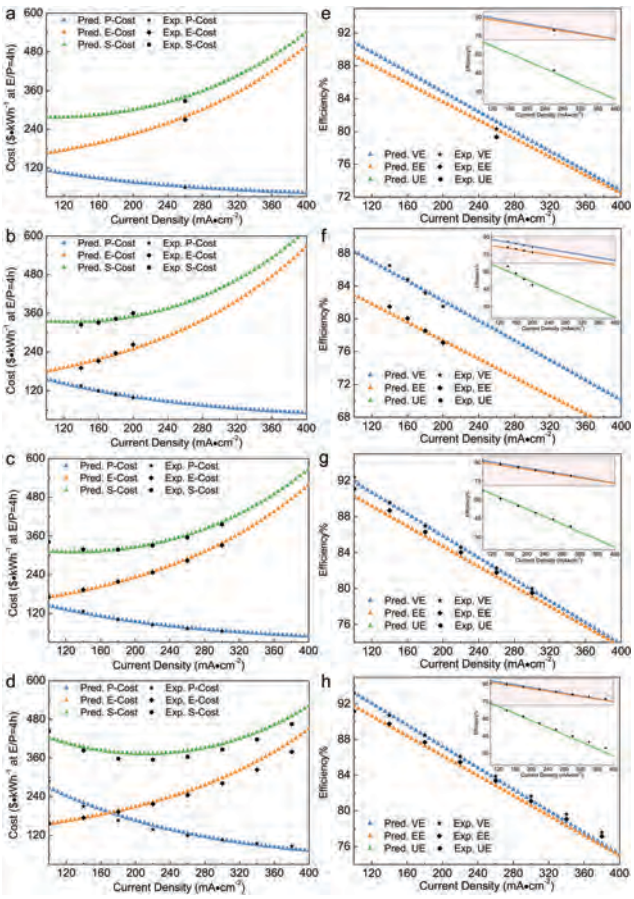

Fig. 2 Comparison of the costs and efficiencies between experiment and the "full features" models. Costs of (a) "VFB 20190423", (b) "VFB 20180920", (c) "VFB 20190225" and (d) "VFB 20200311"; and efficiencies of (e) "VFB 20190423", (f) "VFB 20180920", (g) "VFB 20190225" and (h) "VFB 20200311".

improve the operating current density, which maintains a high power rate of a stack. It can be easily understood that increasing the current density will lead to high polarization and

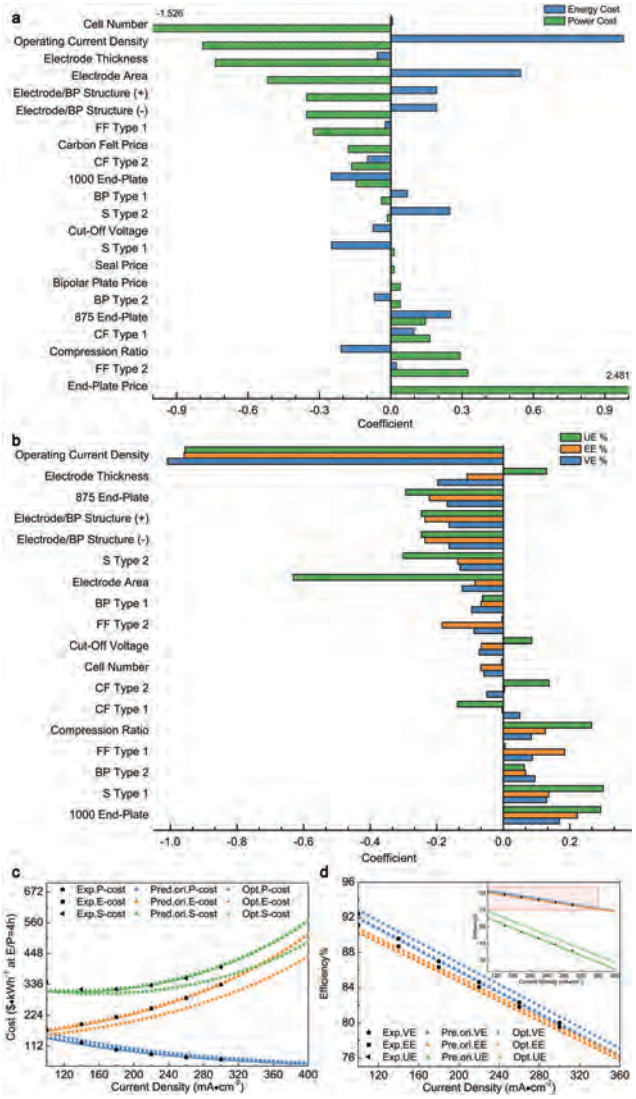

Fig. 3 Coefficients of the "full features" models, (a) P-cost and E-cost; and (b) VE, EE, and UE. Comparison of the cost and efficiencies between experiment and model prediction for the optimized and original stack named "VFB 20190225", (c) costs and (d) efficiencies.

decrease the voltage efficiency, which lowers the utilization of electrolytes. In this case, the P-cost will decrease and the E-cost will increase. Therefore, the S-cost will first decrease and then increase, and a suitable operating current density should be selected.

The stacks are assembled with different kinds of key materials, and their types and properties are listed in Table S3 (ESI†). Apart from exploring key materials and an endplate with a lower cost, the following measures can be adopted to lower the P-cost.

Increase the cell number and electrode area to improve the stack power; decrease the electrode thickness, and lower the contact resistance between the electrode/bipolar plate by optimizing the electrode/bipolar plate structure, using CF Type 2 with higher electrocatalytic activity (see the CV curves of two types of carbon felts in Fig. S5, ESI†) and S Type 2.

Effective ways to lower the E-cost are as follows:

Increase the electrode thickness, cutoff voltage, electrode compression ratio, and electrode/bipolar plate structure, and decrease the operating current density and electrode area to improve the voltage efficiency or decrease the polarization and further improve the UE, using CF Type 2, BP Type 2 and S Type 1.

It can be easily understood that the above methods can affect the P-cost and E-cost due to the following reasons:

(1) Increasing the electrode thickness can provide more reactive sites, which can improve the UE; however, on the other hand a large thickness of the electrodes will cause high ohmic polarization and thus would reduce the EE and VE of a battery.

(2) Increasing the cut-off voltage can increase the charging capacity of the battery, which can improve the UE and thus reduce the E-cost.

(3) Increasing the electrode compression ratio can decrease the electrode thickness, shortening the distance between the current collector and the membrane for ions and electrons, which can decrease the ohmic polarization and thus decrease the E-cost and P-cost.

(4) Using a higher electronic conductivity bipolar plate can lower the ohmic polarization, which can improve the performance of the battery and thus lower the E-cost and P-cost as well.

Some features can simultaneously reduce the P-cost and E-cost, such as using higher electrocatalytic activity carbon felt and a higher electronic conductivity bipolar plate, which should be chosen for the VFB stack design. However, many features are contradictory to reduce the P-cost and E-cost simultaneously, such as the operating current density and electrode area. Thus, there is always an optimal operating current density to balance the P-cost and E-cost for each VFB stack.

From Fig. 3(b), it can be concluded that an effective method to improve the VE and EE is as follows:

Increase the electrode compression ratio; decrease the operating current density, electrode area and thickness; optimize the electrode/bipolar plate structure, cutoff voltage, and cell number; and use FF Type 1 (it has the shortest flow passage and can decrease the concentration polarization) and BP Type 2. Effective methods to improve the UE are like the VE, apart from the electrode thickness and cutoff voltage.

Based on the coefficient of each feature, we optimized one stack named "VFB 20190225" in Database-1 and predicted its efficiencies and cost. A comparison of the predicted results with the experimental data is shown in Fig. 3(c) and (d) and Table S4 (ESI†). The optimal operating current density for the optimized stack increased by 30 mA cm^{-2} (due to using a higher conductivity and low contact resistance bipolar plate) and the UE increased by 2.28%; accordingly, there was a 2.98%, 2.49%, and 2.69% reduction for the P-cost, E-cost, and S-cost, respectively, regardless of the reduction of the VE and EE by 1.03% and 1.42%, respectively. By increasing the cutoff voltage for each cell, the UE increases significantly, so the E-cost reduces accordingly, while the P-cost changes little; thus, the

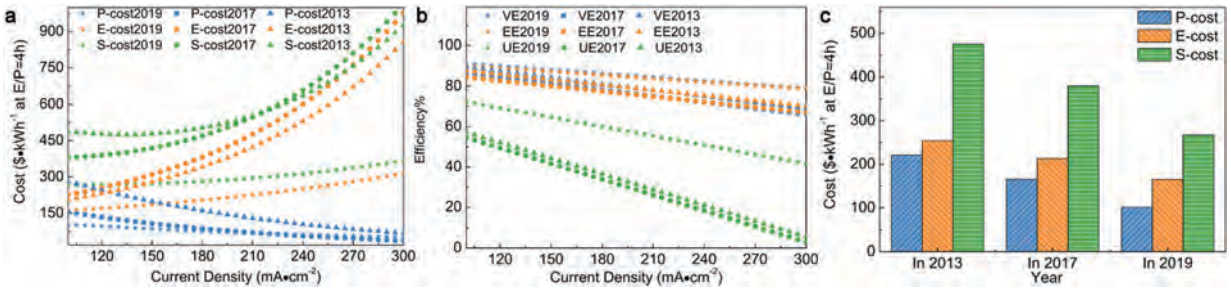

Fig. 4 Predicted results for each stack in 2013, 2017 and 2019. (a) Costs, (b) efficiencies, and (c) comparison of each cost at the optimal operating current density.

S-cost can also be reduced significantly. In summary, future R&D of materials for VFB stacks should aim to reduce their ohmic polarization as previously discussed for vanadium flow batteries with high power density,[53,54] and alter the methods to improve the UE while keeping the P-cost constant. The results confirmed that the ML methodology can effectively clarify the relationship between high-dimensional data such as the design and optimization of VFB stacks, which is likely to accelerate the R&D of VFB stacks.

Comparison of VFB stack efficiencies and costs in 2013, 2017 and 2019

Based on the above ML methodology, which can predict and optimize the efficiencies and cost of VFB stacks, another VFB dataset (Dataset-2) was established by using the data from the years of 2012 to 2018. The "full features" models were still applied to predict the efficiencies and costs for Dataset-2. The precision analysis of the ML model to predict the efficiencies and cost of VFB stacks in Dataset-2 is shown in Table S5 (ESI†). The operating current density range is 80 mA cm^{-2} to 200 mA cm^{-2} in Dataset-2. To have a comprehensive comparison with our recent generation VFB stacks, we chose 3 stacks that were assembled and tested in 2013, 2017 and 2019, respectively.

With the help of the ML methodology, the predicted VE, EE, UE, P-cost, E-cost and S-cost of these stacks working at different current densities can be compared and are shown in Fig. 4(a)–(c) and Table 2. Overall, the S-cost under the optimal operating current density is reduced from \$475.36 per kW h (at E/P = 4 h) in 2013 to \$379.83 per kW h (at E/P = 4 h) in 2017 and then to \$267.21 per kW h (at E/P = 4 h) in 2019. The efficiencies of the stack in 2017 are slightly less than those in 2013. However, because the use of CF Type 1 is vastly cheaper than that of CF Type 2, the optimal operating current density changed from 135 mA cm^{-2} to 90 mA cm^{-2}, which led to higher efficiencies and a lower cost. With the optimization of the VFB stack structure, the efficiencies in 2019 are the highest with the same electrode material CF Type 1, which led to a further cost reduction. The optimal operating current density changed from 90 mA cm^{-2} in 2017 to 110 mA cm^{-2} in 2019, while the efficiencies, especially the UE, improved. This result guides future R&D of VFB stacks to improve the utilization efficiency of electrolytes under high operating current densities.

In recent years, we have tried to decrease the polarization of the stack by optimizing the battery structure (flow field, electrode inlet diversion structure, compression ratio, *etc.*), developing high performance key materials *e.g.* exploring membranes with improved selectivity and conductivity, electrodes with higher electrochemical activity, and bipolar plates with higher electron conductivity. As a result, the power density was increased dramatically and further the cost was reduced.

Analysis of the VFB system cost proportion with different cell numbers

Based on the materials and structure optimization of the stack named "VFB 20190225," we changed the cell number to 5 or 30 (other parameters will also change with the value of the cell number) and predicted the cost for each stack with an operating current density range from 100 mA cm^{-2} to 400 mA cm^{-2} [Fig. 5(a)]. Then, the cost proportion of each part of the three stacks would vary with the operating current density accordingly [Fig. 5(b)–(d)]. With the increase in cell number, there was a remarkable reduction of the P-cost without a marked change of the E-cost, which led to a notable reduction of the S-cost. The electrolyte cost accounts for approximately 60% at the

Table 2 Comparison of efficiencies and costs between three stacks assembled in three different years at the optimal operating current density

	Stack in 2013	Stack in 2017	Stack in 2019
The optimal operating current density (mA cm^{-2})	135	90	110
VE (%)	85.69	87.31	90.40
EE (%)	84.52	85.32	89.32
UE (%)	48.20	57.17	70.70
P-cost	221.13	165.99	101.53
E-cost	254.24	213.85	165.69
S-cost	475.36	379.83	267.21

The unit of P-cost and E-cost is (\$ (kW h)$^{-1}$ at E/P = 4 h).

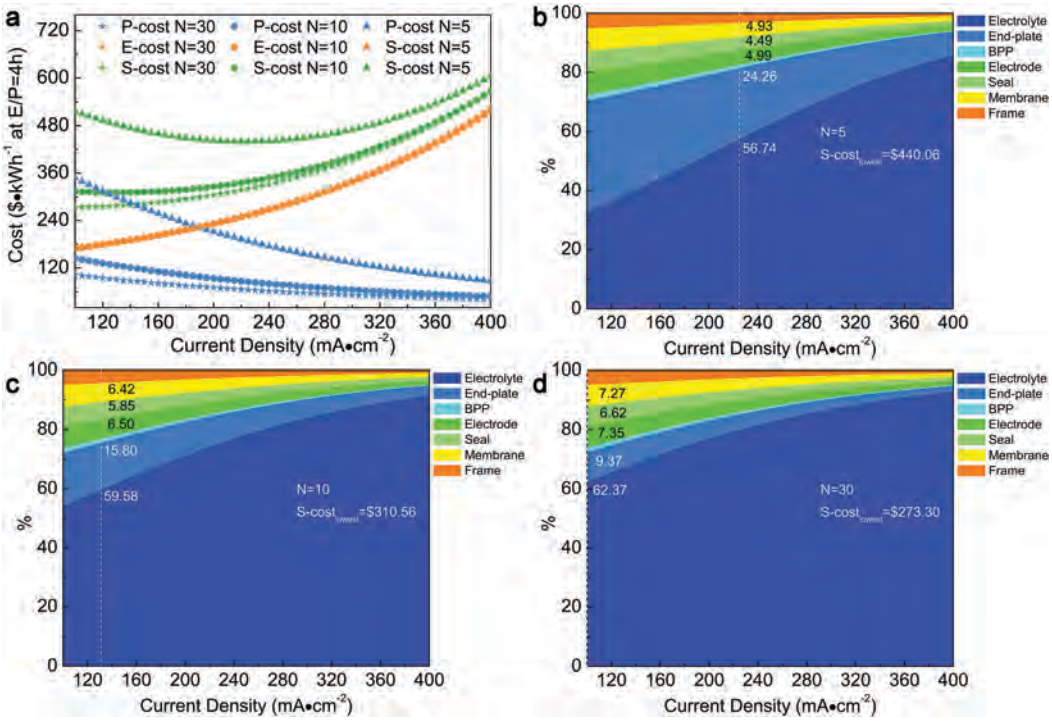

Fig. 5 (a) Comparison of the costs of the stacks with the same structure but different cell number; and the change of the proportion of the key materials in the VFB system with the operating current density with (b) cell number = 5, (c) cell number = 10, and (d) cell number = 30. The white dotted line in (b–d) represents the optimal operating current density of each stack.

optimal operating current density for each stack, which has the highest share, and its proportion increased significantly with the increase in operating current density. Therefore, the most effective way to reduce the S-cost is to reduce the E-cost (improve the UE) at high operating current density.

Cost change of a VFB system with different VE and UE

Finally, the relationship between the efficiencies and S-cost for our optimized VFB stack of the latest generation was predicted based on the "full features" models and current material costs. The mappings of current density (mA cm^{-2})–VE (%)–S-cost ($ (kW h)$^{-1}$ at E/P = 4 h) with a UE range from 60% to 100% are shown in Fig. 6(a)–(i). For example, if the current density is 200 mA cm^{-2}, and the VE of a stack reaches 90%, when the UE is maintained at 60%, the S-cost is about $273 per kW h; if the UE increases to 80%, the S-cost will decrease to around $223 per kW h (E/P = 4 h). However, the cost can further decrease to $206 per kW h (E/P = 4 h) when UE increases to 90%. To further optimize the operating conditions and reduce the cost, the S-cost at E/P = 10 h and E/P = 20 h was predicted and the results are shown in Fig. S4 (ESI†). If the UE is maintained at 80%, the S-cost is about $182 per kW h at E/P = 10 h and $165 per kW h at E/P = 20 h, which are 18% and 26% lower than that at E/P = 4 h. The results indicate that a VFB is more suitable for longer duration energy storage. In summary, the future R&D of VFB stacks should still focus on high current density with both high VE and UE to rapidly promote their industrialization in large-scale energy storage.

To realize the above target, the work on key materials and the stack architecture needs to be focused to decrease the polarization and further improve the power density and UE in the future. For example, to explore membranes with higher conductivity, electrodes with higher activity and conductivity are preferred to decrease the ohmic and electrochemical polarization. Further optimization of the battery structure is also very important to decrease the concentration polarization and to improve the battery performance and further decrease the cost.

Conclusions

In summary, we successfully integrated the state-of-art ML methodology with our experiments in the R&D of VFB stacks, and optimized their materials and structures, and then predicted future development. This methodology could remarkably shorten the R&D and industrialization of VFB technology and it can be extended to other materials and structure design or the equipment-operating condition optimization of flow batteries. The models can reach a prediction accuracy of MAPE within 1% for both the VE and EE between experiment and predicted results, and within 5.2% for the UE, P-cost and E-cost, which is regarded as reliable and can be used to assist the optimization of VFB stacks and predict their performance and cost. As the E-cost accounts for the highest proportion (about 60% at the optimal operating current density) of the total cost within a VFB system, the core objective is to improve the UE of VFB stacks at a high current density, which would

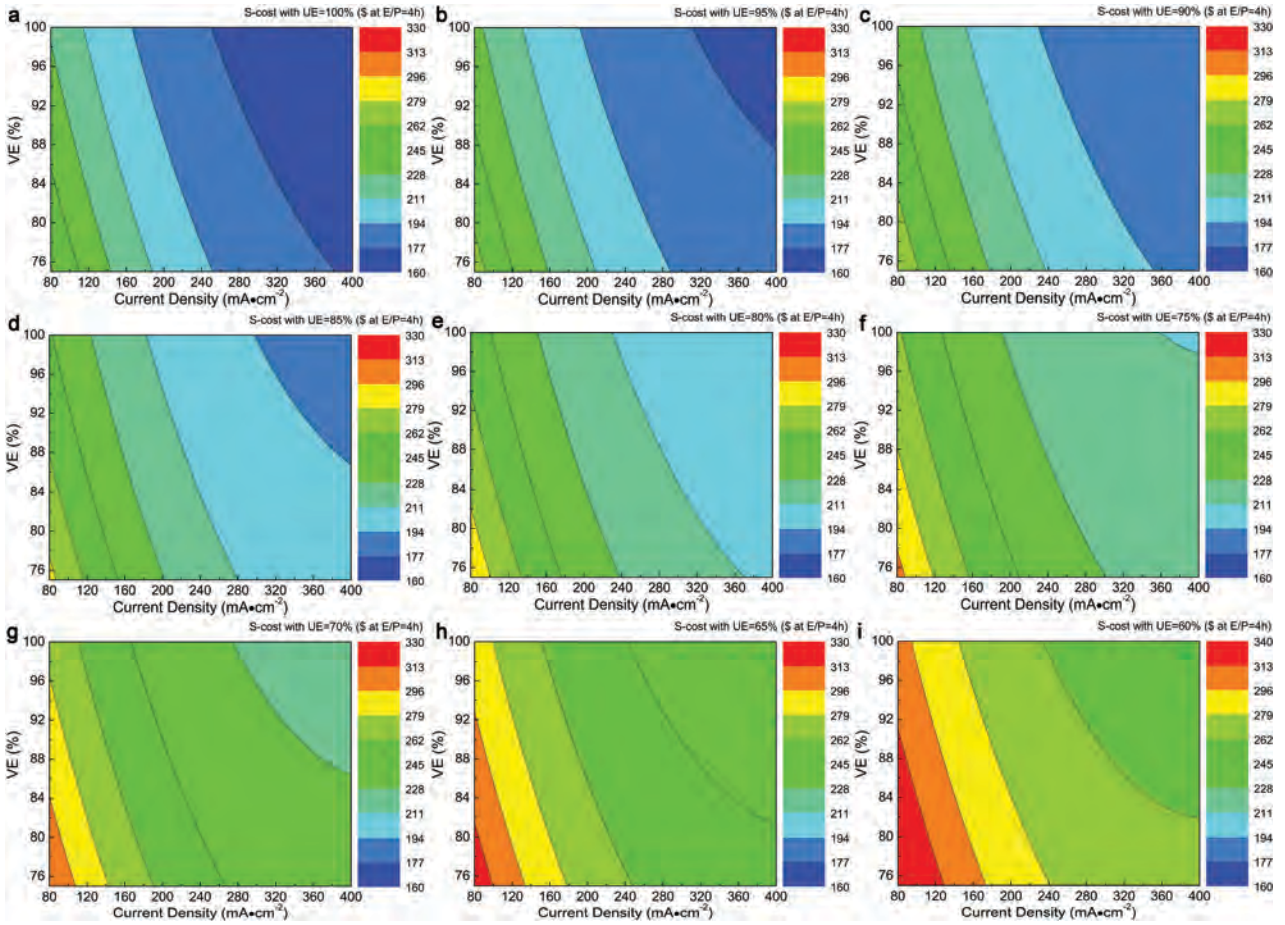

Fig. 6 Prediction of the S-cost change with the operating current density at different VE and UE (E/P = 4 h). (a) UE = 100%, (b) UE = 95%, (c) UE = 90%, (d) UE = 85%, (e) UE = 80%, (f) UE = 75%, (g) UE = 70%, (h) UE = 65%, and (i) UE = 60%.

promote their rapid industrialization in large-scale energy storage. For example, the development of materials for the VFB stack should focus on reducing the electrochemical polarization and ohmic polarization at high current densities, and the design of the flow field should monitor the enhancement of mass transfer to decrease the concentration polarization of the flow battery stacks.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (21925804, U1808209, 21706253), Strategic Priority Research Program of the CAS (XDA21070100), CAS Engineering Laboratory for Electrochemical Energy Storage and Liaoning Revitalization Talents Program (XLYC1802050). We thank Zhizhang Yuan, Yanbin Yin, Meng Yue and Qiong Zheng for discussions. We thank Shijie Zhu and Bin Hao for helping us draw part of the Table of Content.

References

1 D. Larcher and J. M. Tarascon, *Nat. Chem.*, 2015, **7**, 19–29.
2 B. Dunn, H. Kamath and J.-M. Tarascon, *Science*, 2011, **334**, 928–935.
3 P. Leung, X. Li, C. Ponce de León, L. Berlouis, C. T. J. Low and F. C. Walsh, *RSC Adv.*, 2012, **2**, 10125–10156.
4 H. Chen, T. N. Cong, W. Yang, C. Tan, Y. Li and Y. Ding, *Prog. Nat. Sci.*, 2009, **19**, 291–312.
5 Q. Dai, Z. Liu, L. Huang, C. Wang, Y. Zhao, Q. Fu, A. Zheng, H. Zhang and X. Li, *Nat. Commun.*, 2020, **11**, 13.
6 H. Zhang, H. Zhang, X. Li, Z. Mai and J. Zhang, *Energy Environ. Sci.*, 2011, **4**, 1676–1679.
7 H. Zhang, H. Zhang, X. Li, Z. Mai and W. Wei, *Energy Environ. Sci.*, 2012, **5**, 6299–6303.
8 H. Zhang, H. Zhang, F. Zhang, X. Li, Y. Li and I. Vankelecom, *Energy Environ. Sci.*, 2013, **6**, 776–781.
9 W. Lu, D. Shi, H. Zhang and X. Li, *Energy Storage Mater.*, 2019, **17**, 325–333.
10 L. Qiao, H. Zhang, W. Lu, Q. Dai and X. Li, *ACS Appl. Mater. Interfaces*, 2019, **11**, 24107–24113.
11 Z. Yuan, Y. Duan, H. Zhang, X. Li, H. Zhang and I. Vankelecom, *Energy Environ. Sci.*, 2016, **9**, 441–447.

12 Y. Zhao, W. Lu, Z. Yuan, L. Qiao, X. Li and H. Zhang, *J. Mater. Chem. A*, 2017, **5**, 6193–6199.
13 M. Park, J. Ryu, W. Wang and J. Cho, *Nat. Rev. Mater.*, 2016, **2**, 16080.
14 X. Zhang, Q. Wu, Y. Lv, Y. Li and X. Zhou, *J. Mater. Chem. A*, 2019, **7**, 25132–25141.
15 A. Mukhopadhyay, Y. Yang, Y. Li, Y. Chen, H. Li, A. Natan, Y. Liu, D. Cao and H. Zhu, *Adv. Funct. Mater.*, 2019, **29**, 1903192.
16 T. Liu, X. Li, H. Nie, C. Xu and H. Zhang, *J. Power Sources*, 2015, **286**, 73–81.
17 M. Jiao, T. Liu, C. Chen, M. Yue, G. Pastel, Y. Yao, H. Xie, W. Gan, A. Gong, X. Li and L. Hu, *Energy Storage Mater.*, 2020, **27**, 327–332.
18 W. Liao, F. Jiang, Y. Zhang, X. Zhou and Z. He, *Renewable Energy*, 2020, **152**, 1310–1316.
19 K. Yaji, S. Yamasaki, S. Tsushima, T. Suzuki and K. Fujita, *Struct. Multidiscipl. Optim.*, 2018, **57**, 535–546.
20 M. Yue, Q. Zheng, F. Xing, H. Zhang, X. Li and X. Ma, *AIChE J.*, 2018, **64**, 782–795.
21 Q. Zheng, F. Xing, X. Li, G. Ning and H. Zhang, *J. Power Sources*, 2016, **324**, 402–411.
22 Q. Xu, T. S. Zhao and P. K. Leung, *Appl. Energy*, 2013, **105**, 47–56.
23 X. Wu, X. Yuan, Z. Wang, J. Liu, Y. Hu, Q. Deng, X. Yin, Q. Zhou, W. Zhou and Y. Wu, *J. Solid State Electrochem.*, 2017, **21**, 429–435.
24 P. Zhao, H. Zhang, H. Zhou, J. Chen, S. Gao and B. Yi, *J. Power Sources*, 2006, **162**, 1416–1420.
25 C.-H. Chen, K. Yaji, S. Yamasaki, S. Tsushima and K. Fujita, *J. Energy Storage*, 2019, **26**, 100990.
26 D. K. Kim, S. J. Yoon, J. Lee and S. Kim, *Appl. Energy*, 2018, **228**, 891–901.
27 M. Yue, Z. Lv, Q. Zheng, X. Li and H. Zhang, *Appl. Energy*, 2019, **235**, 495–508.
28 T. Liu, X. Li, H. Zhang and J. Chen, *J. Energy Chem.*, 2018, **27**, 1292–1303.
29 A. Bhattacharjee and H. Saha, *J. Energy Storage*, 2017, **13**, 220–232.
30 M. Guarnieri, A. Trovò, G. Marini, A. Sutto and P. Alotto, *J. Power Sources*, 2019, **431**, 239–249.
31 Q. Zheng, X. Li, Y. Cheng, G. Ning, F. Xing and H. Zhang, *Appl. Energy*, 2014, **132**, 254–266.
32 C. Minke, U. Kunz and T. Turek, *J. Power Sources*, 2017, **342**, 116–124.
33 C. Minke and T. Turek, *J. Power Sources*, 2015, **286**, 247–257.
34 C. Minke, U. Kunz and T. Turek, *J. Power Sources*, 2017, **361**, 105–114.
35 J. Noack, L. Wietschel, N. Roznyatovskaya, K. Pinkwart and J. Tuebke, *Energies*, 2016, **9**, 627.
36 M.-J. Li, W. Zhao, X. Chen and W.-Q. Tao, *Appl. Therm. Eng.*, 2017, **114**, 802–814.
37 P. Raccuglia, K. C. Elbert, P. D. F. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier and A. J. Norquist, *Nature*, 2016, **533**, 73–76.
38 K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev and A. Walsh, *Nature*, 2018, **559**, 547–555.
39 N. C. Frey, J. Wang, G. I. Vega Bellido, B. Anasori, Y. Gogotsi and V. B. Shenoy, *ACS Nano*, 2019, **13**, 3031–3041.
40 Z. Hou, Y. Takagiwa, Y. Shinohara, Y. Xu and K. Tsuda, *ACS Appl. Mater. Interfaces*, 2019, **11**, 11545–11554.
41 G. H. Gu, J. Noh, I. Kim and Y. Jung, *J. Mater. Chem. A*, 2019, **7**, 17096–17117.
42 M. H. S. Segler, M. Preuss and M. P. Waller, *Nature*, 2018, **555**, 604–610.
43 J. M. Granda, L. Donina, V. Dragone, D.-L. Long and L. Cronin, *Nature*, 2018, **559**, 377–381.
44 P. S. Lamoureux, K. T. Winther, J. A. G. Torres, V. Streibel, M. Zhao, M. Bajdich, F. Abild-Pedersen and T. Bligaard, *ChemCatChem*, 2019, **11**, 3579–3599.
45 Y. Amar, A. Schweidtmann, P. Deutsch, L. Cao and A. Lapkin, *Chem. Sci.*, 2019, **10**, 6697–6706.
46 J. Timoshenko and A. I. Frenkel, *ACS Catal.*, 2019, **9**, 10192–10211.
47 A. Smith, A. Keane, J. A. Dumesic, G. W. Huber and V. M. Zavala, *Appl. Catal., B*, 2020, **263**, 118257.
48 M. Sun, A. W. Dougherty, B. Huang, Y. Li and C.-H. Yan, *Adv. Energy Mater.*, 2020, **10**, 1903949.
49 A. Mosavi, S. Shamshirband, E. Salwana, K.-W. Chau and J. H. M. Tah, *Eng. Appl. Comput. Fluid Mech.*, 2019, **13**, 482–492.
50 A. Baghban, M. Kahani, M. A. Nazari, M. H. Ahmadi and W.-M. Yan, *Int. J. Heat Mass Transfer*, 2019, **128**, 825–835.
51 K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis, M. Z. Bazant, S. J. Harris, W. C. Chueh and R. D. Braatz, *Nat. Energy*, 2019, **4**, 383–391.
52 P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y.-H. Liao, M. H. Chen, B. Cheong, N. Perkins, Z. Yang, P. K. Herring, M. Aykol, S. J. Harris, R. D. Braatz, S. Ermon and W. C. Chueh, *Nature*, 2020, **578**, 397–402.
53 C. Ding, H. Zhang, X. Li, T. Liu and F. Xing, *J. Phys. Chem. Lett.*, 2013, **4**, 1281–1294.
54 W. Lu, X. Li and H. Zhang, *Phys. Chem. Chem. Phys.*, 2017, **20**, 23–35.