Article:
Liu, K, Wei, Z, Yang, Z et al. (2021) Mass load prediction for lithium-ion battery electrode clean production: a machine learning approach. Journal of Cleaner Production, 289. 125159. ISSN: 0959-6526

https://doi.org/10.1016/j.jclepro.2020.125159

Mass load prediction for lithium-ion battery electrode clean production：a machine learning approach

Kailong Liu[a]*, Zhongbao Wei[b], Zhile Yang[c], Kang Li[d]*

a WMG, The University of Warwick, Coventry, CV4 7AL, United Kingdom {Email: kliu02@qub.ac.uk, kailong.liu@warwick.ac.uk}.

b National Engineering Laboratory for Electric Vehicles, School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China {Email: weizb@bit.edu.cn}.

c Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, 518055, China {Email: zl.yang@siat.ac.cn}.

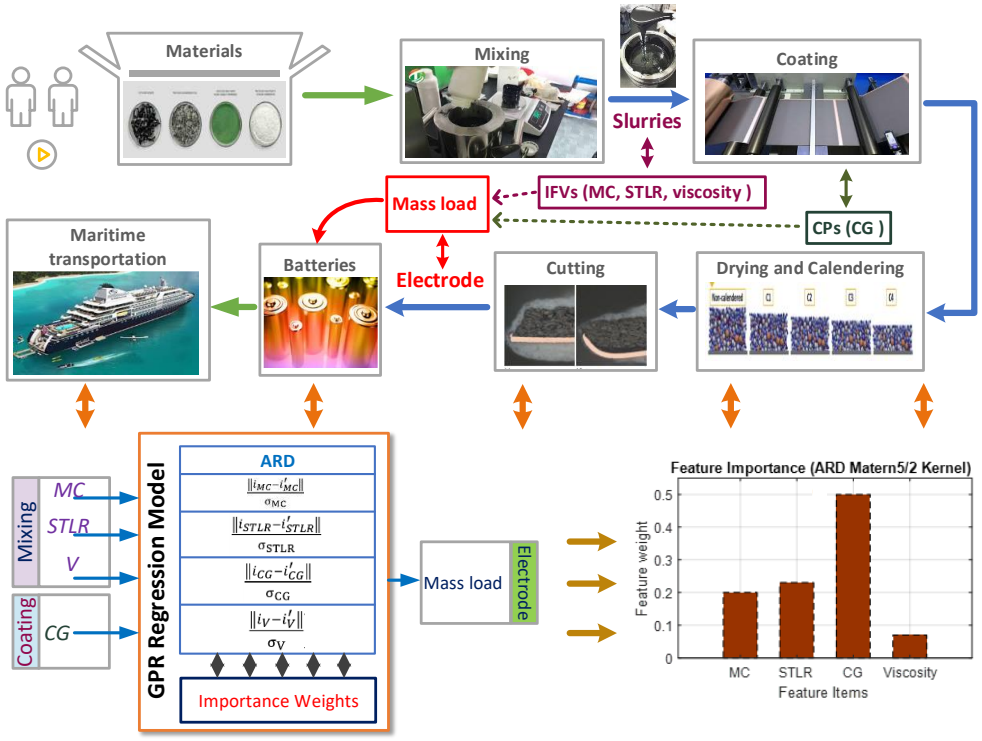
d School of Electronic and Electric Engineering, University of Leeds, LS2 9JT, United Kingdom {Email: k.li1@leeds.ac.uk}

** Corresponding author*

Abstract: With the advent of sustainable and clean energy, lithium-ion batteries have been widely utilised in cleaner productions such as energy storage systems and electrical vehicles, but the management of their electrode production chain has a direct and crucial impact on the battery performance and production efficiency. To achieve a cleaner production chain of battery electrode involving strongly-coupled intermediate parameters and control parameters, a reliable approach to quantify the feature importance and select the key feature variables for predicting battery intermediate products is urgently required. In this paper, a Gaussian process regression-based machine learning framework, which incorporates powerful automatic relevance determination kernels, is proposed for directly quantifying the importance of four intermediate production feature variables and analysing their influences on the prediction of battery electrode mass load. Specifically, these features include three intermediate parameters from the mixing step and a control parameter from the coating step. After deriving four different automatic relevance determination kernels, the importance of these four feature variables based on a regression modelling is comprehensively analysed. Comparative results demonstrate that the proposed automatic relevance determination kernel-based Gaussian process regression models could not only quantify the importance weights for reliable feature selections but also help to achieve satisfactory electrode mass load prediction. Due to the data-driven nature, the proposed framework can be conveniently extended to improve the analysis and control of battery electrode production, further benefitting the manufactured battery yield, efficiencies and performance to achieve cleaner battery production.

Key words: Battery electrode production; data-driven model; mass load prediction; efficient energy storage system; cleaner production.

Graphical Abstract:



A list of terminology and symbols

Li-ion	lithium-ion
EV	electrical vehicle
CRISP	cross-industry standard process
IFV	intermediate feature variable
CP	control parameter
GPR	Gaussian process regression
ARD	automatic relevance determination
EX	exponential
SE	squared-exponential
NMP	N-Methyl-2-pyrrolidone
PVDF	polyvinylidene difluoride
MC	mass content
STLR	solid-to-liquid ratio
CG	comma gap
MAE	mean absolute error
RMSE	root mean square error
$m(j)$	mean function
$k(j,j')$	covariance function
θ_{hp}	hyper-parameters
$L(\theta_{hp})$	negative log marginal likelihood
$p(y'\|j,y,j')$	conditional distributions

$\bar{y}'$	mean prediction value
$\boldsymbol{N}$	distribution matrix
$\boldsymbol{cov}$	variance matrix
k_{EX}	exponential kernel
k_{SE}	squared-exponential kernel
$k_{M3/2}$	Matern3/2 kernel
$k_{M5/2}$	Matern5/2 kernel
k_{ARDEX}	exponential kernel with ARD structure
k_{ARDSE}	squared-exponential kernel with ARD structure
$k_{ARDM3/2}$	Matern3/2 kernel with ARD structure
$k_{ARDM5/2}$	Matern5/2 kernel with ARD structure
W	feature weights

1. Introduction

Nowadays, lithium-ion (Li-ion) batteries have been widely utilised to boost the development of cleaner productions such as electrical vehicles (EVs) and energy storage systems, due to their low discharge-rates and high energy densities (Liu, C. et al., 2019). However, the performance of Li-ion batteries would be directly and highly affected by their electrode production processes, which significantly hinders the improvements of battery technology. The battery electrode production is of great concern in developing clean and effective energy storage systems, which is a key factor in securing tangible economic payback and in improving the efficiencies of large-scale clean energy applications (Dahodwalla and Herat, 2000). In light of this, efforts are urgently needed to fully understand the intermediate products and parameters within the battery electrode production chain (He et al., 2020).

Unfortunately, battery electrode production chain contains numerous intermediate processes with a large amount of intermediate products, parameters and influencing factors (Lee et al., 2010). Due to the complexity and strong-coupled interdependencies of electrode manufacturing steps, the multiple correlations among feature variables of intermediate products and control parameters are still difficult to model. As the whole battery production chain consists of a number of chemical, mechanical as well as electrical operations and would generate over 600 influencing parameters or variables, the analysis of feature variables in electrode production requires the deep expert experiences and specialized equipment, which still mainly relies on the trial and error solutions (Kwade et al., 2018; Pang et al., 2019). Therefore, in order to achieve smarter and cleaner battery production, advanced data analysis strategies to better quantify feature variables and select key feature items for predicting battery electrode properties are urgently needed.

With the rapid development of artificial intelligence and machine learning technologies, data-driven strategies have become popular in the field of battery management (Li et al., 2019; Lipu et al., 2018; Zhou et al., 2020). A range of data-driven approaches have been designed for estimating battery states (Meng et al., 2019; Tian et al., 2020; Wei et al., 2019; Zhou et al., 2019), predicting battery lifetime during cycling (Tang et al., 2019; Zhang et al., 2019) or storage modes (Liu et al., 2019), diagnosing battery faults (Pan et al., 2020; Yang, R. et al., 2018), balancing battery cells (Ouyang et al., 2020; Saw et al., 2016), achieving efficient thermal management (Shang et al., 2019; Xiong et al., 2018) and charging management (Liu et al., 2018; Maia et al., 2019). Overall, through designing suitable data-driven models, it is expected that smarter and more efficient management of Li-ion batteries can be achieved. However, these researches primarily focus on the battery macroscopic performance without taking the battery intermediate properties in the production process into account. It should be noted that the battery production chain also generates a large amount of data and plays a more direct role in determining the battery performance, designing an effective data-driven approach to quantify and predict battery intermediate features is therefore also crucial for boosting the development of cleaner production (Dahodwalla and Herat, 2000).

In contrast to the battery management where fruitful solutions are available, less reports have been found so far on using advanced machine learning technologies to improve the battery production (Wanner et al., 2019). Among limited literatures on battery productions (e.g. monitoring (Knoche et al., 2016), adjustment (Schünemann et al., 2016) and control (Günther et al., 2020)), developing a proper data-driven based model to analyse feature variables and predict intermediate product properties is a hotspot. For instance, a data-driven approach was proposed in (Schnell and Reinhart, 2016) to analyse the failure modes and parametric effects, which contributes to the improvement of battery production chain control. Based on the cross-industry standard process (CRISP), Schnell et al. (Schnell et al., 2019) designed a linear model and a neural network model to identify the process dependencies and forecast the battery production properties. Turetskyy et al. (Turetskyy et al., 2019) utilised the decision tree techniques to analyse feature importance and forecast the maximum capacity of battery. A multi-variate data-driven model was designed in (Thiede et al., 2019) to discover proper quality gates for predicting the manufactured battery properties. Based upon the statistical analyses of fluctuations in battery production, the influence of these fluctuations on manufactured battery capacity is evaluated in (Hoffmann et al., 2019). In (Cunha et al., 2020), several 2D graphs produced by three conventional machine learning classification models are used to analyse the dependencies of battery production features. Despite the aforementioned works on the data-driven

modelling of battery production, several limitations and challenges remain to be improved: 1) most researches simply apply conventional methodologies to predict the properties of battery production, little has been done to comprehensively evaluate the data-driven approaches to further improve their performance in battery production area. 2) commonly-used machine learning technologies such as support vector machines and neural networks only produce predictions of battery production properties, but fail to directly quantify the importance of battery production features which is also of interests to engineers. Such information could benefit engineers to conduct sensitivity analysis for effective feature selection and optimise the production chain to further enhance the manufactured batteries performance and achieve cleaner battery production.

Given the aforementioned considerations, this article deals with the data-driven electrode property prediction for Li-ion battery production, where the importance weights of multiple battery production feature items can be also taken into account. The main contributions of this paper are summarized as follows: 1) After identifying four feature items of interest from three intermediate feature variables (IFVs) in the mixing step and a control parameter (CP) in the coating step of the battery production process, a Gaussian process regression (GPR)-based machine learning framework is designed to effectively map the underlying relations among these feature items and the battery electrode mass load. 2) The kernel functions within GPR are enhanced with an automatic relevance determination (ARD) structure, allowing multiple input variables to obtain different scales of hyper-parameters to improve the regression accuracy and robustness. 3) Based upon the well-tuned hyper-parameters within the ARD kernels, the importance weights of these feature variables of interest could be effectively quantified for the reliable selections of battery production features. 4) The importance of these feature items and the prediction results are comprehensively evaluated and quantified with four different ARD kernels derived from the exponential (EX), squared-exponential (SE), Matern3/2 and Matern5/2 covariance functions. This is the first known application by designing the ARD kernel-based GPR to directly and simultaneously quantify the weights of battery feature items and analyse their effects on the regressions of battery electrode mass load. Obviously, the proposed ARD kernel-based GPR framework could bring the significant benefits such as better monitoring the key variables within battery production chain to enhance the manufactured battery yield, efficiencies and performance, as well as saving the costs of battery production, further leading to cleaner battery production and improving the optimisation of energy storage system.

The remainder of this paper is organized as follows. Section 2 describes some key steps and variables within the battery electrode production chain. Section 3 presents the fundamental of GPR, ARD kernels, detailed feature selection workflow and several performance indicators. Section 4 gives an in-depth comparison of the prediction results of electrode mass load and discuss the quantified importance weights of all features of interest. Finally, Section 5 concludes this paper.

2. Battery electrode production chain

Electrode production is a primary and significantly sophiscated process in the battery manufacturing chain, covering multiple disciplines such as mechanical engineering, chemical engineering and electrical engineering (Chouchane et al., 2019). Moreover, the properties of manufactured electrode such as mass load plays a key role in determining the battery final performance (energy or power densities, maximal capacity and service life), which must be carefully managed for benefitting the cleaner battery production.

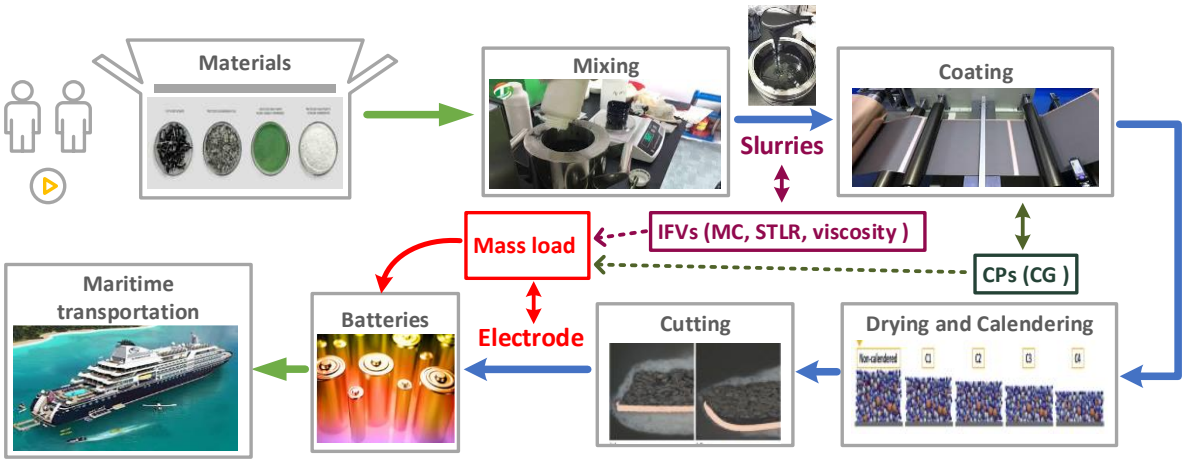


Fig. 1. A systematic framework of some key steps in battery electrode production chain.

Fig. 1 gives a systematic framework illustrating some key steps in the battery electrode production chain. It should be known that battery electrode manufacturing is a lengthy process involving many intermediate and individual steps (i.e., mixing, coating, drying, calendaring and cutting) as well as specialized equipment (i.e., mixer, coater, oven). In this framework, the first key process is the mixing step which produces the slurries for both anodes and cathodes. Specifically, in the mixing step, the prepared active materials such as graphite and Li-NCM-Oxide, the conductive additive material such as carbon black, the solvent such as N-Methyl-2-pyrrolidone (NMP) as well as the binder such as polyvinylidene difluoride (PVDF) will be mixed within a soft blender. A standard process is

that the binder is first dissolved into the solvent, followed by adding the active and conductive additive materials to produce slurries. After mixing, these mixed slurries would be coated upon both two surface sides of mental foils by a comma-gap coater. In general, the copper foil is utilised for anode coating while the aluminium foil is adopted for cathode coating. In this stage, the speed of coater would be fixed and the comma gap is changed to ensure shear force for determining the coating thickness. Moreover, the wet coating products will be dried by setting constant temperatures of several built-in ovens simultaneously. Then, a calendaring step will follow to evaporate the residual solvent of dried coating products. Finally, both anode and cathode electrodes are obtained after cutting the coating products into proper sizes. It should be known that numerous variables and parameters are involved in such a complex production chain. Several IFVs and CPs are of extreme importance for determining the properties of electrodes, which would further affect the battery production results and must be well studied and controlled.

In this context, three key IFVs including the active material's mass content (MC), solid-to-liquid ratio (STLR) and slurry viscosity from mixing process as well as one CP named coater comma gap (CG) from coating process are selected as the interested feature variables. Then an ARD kernel-based GPR framework is derived to quantify the weights of these four interested features and analyse their effects on the prediction performance of battery electrode mass load. In theory, STLR could reflect the mass ratios among the solid components (i.e., active materials, conductive additive materials as well as binder) and the mass of slurries (i.e., solvents and solid components). Slurry viscosity would highly affect the coating step and is generally related to the shear rates of coating. CG reflects the gap value between the comma roll and coating roll, which could impact the weight and thickness of coating products. To describe battery electrode property, the mass load of uncalendered electrode with the unit of mg/cm^2 is utilised because it has a direct relation with the final electrode capacity. To well analyse these feature variables and select the most important features for predicting battery electrode property, the original dataset from Franco Laboratoire de Reactivite et Chimie des Solides (LRCS) is used in this study. The effectiveness of this dataset has been validated in (Cunha et al., 2020). More detailed information regarding the experimental design and process description can be found in (Cunha et al., 2020), which would not be repeatedly described here due to space limitation. According to this dataset, effective ARD-kernel based GPR models can be developed for predicting the electrode mass load. The feature weights as well as regression performance of various ARD kernels can be also quantified.

3. Technology

In this section, the fundamental of GPR is first introduced, followed by the descriptions of designing various ARD kernels within GPR for feature weight analysis and selection. Furthermore, the performance indicators to quantify and evaluate model regression performance are also provided.

3.1 Gaussian process regression

Based upon the Bayesian framework, GPR provides a random procedure to conduct non-parametric regressions by the Gaussian process (Liu, K. et al., 2019; Tagade et al., 2020). For any inputs, GPR generally contains a mean function $m(j)$ and a covariance function $k(j,j')$ to reflect the probability distributions as follows:

$$f(j)\sim GPR\big(m(j),k(j,j')\big) \tag{1}$$

where $m(j)$ and $k(j,j')$ can be further expressed as:

$$\begin{cases} m(j)=E(f(j)) \\ k(j,j')=E\big[\big(m(j)-f(j')\big)\big(m(j)-f(j')\big)\big] \end{cases} \tag{2}$$

In real applications, $m(j)$ is usually set as zero to simplify the computational effort (Liu et al., 2020). Covariance function $k(j,j')$ is also called as the kernel function to reflect the correlations between the target observations of training dataset and the prediction outputs by the similarities of respective input.

For a regression, the prior distributions of output could be defined as:

$$y\sim \boldsymbol{N}\big(0,k(j,j')\big) \tag{3}$$

where $\boldsymbol{N}(.)$ is a distribution matrix. Supposing both training set j and testing set j' have the same Gaussian distribution, then the regression test outputs y' will provide the joint prior distributions with the training outputs y as (Rasmussen and Williams, 2006):

$$\begin{bmatrix} y \\ y' \end{bmatrix}\sim \boldsymbol{N}\left(0,\begin{bmatrix} k(j,j) & k(j,j') \\ k(j,j')^T & k(j',j') \end{bmatrix}\right) \tag{4}$$

where $k(j,j), k(j',j')$ are the covariance values among inputs of training set and testing set, respectively. $k(j,j')$ stands for the covariance values among inputs from training and testing sets.

To achieve an efficient regression based on GPR model, the hyper-parameters θ_{hp} within covariance or kernel function should be well tuned during training stage. A common and effective tuning way is through using gradient

descent or heuristic methods to minimize the negative log marginal likelihood $\mathrm{L}(\theta_{hp})$ (Richardson et al., 2019; Yang, D. et al., 2018) as:

$$\mathrm{L}(\theta_{hp}) = -logp(\mathrm{y}|\theta_{hp}) = \frac{1}{2}log[det\ k(\theta_{hp})] + \frac{1}{2}y^T k^{-1}(\theta_{hp})y + \frac{N}{2}log(2\pi) \quad (5)$$

After tuning all θ_{hp} within GPR model, the forested outputs y' related to the dataset j' could be obtained by computing its corresponding conditional distributions $p(y'|j, y, j')$ by:

$$p(y'|j, y, j') \sim \boldsymbol{N}(y'|\bar{y}', cov(y')) \quad (6)$$

with

$$\begin{cases} \bar{y}' = k(i, i')^T [k(i, i)]^{-1} y \\ \boldsymbol{cov}(y') = k(i', i') - k(i, i')^T [k(i, i)]^{-1} k(i, i') \end{cases} \quad (7)$$

where $\bar{y}'$ reflects the mean value of predictions. $\boldsymbol{cov}(y')$ is a variance matrix of these predictions.

3.2 ARD kernel structures

According to the discussions in the above subsection, the kernel function within GPR must be carefully designed because it significantly affects the performance of GPR. Although there exist several powerful kernels for GPR in the literature, a suitable one designed for a specific application is favourable (Zhang et al., 2020). For real regression applications of GPR, some typical and efficient kernels are particularly noteworthy.

The first one is the EX kernel with an exponential form as:

$$k_{EX}(i, i') = \sigma_{\mathrm{EX}}^2 exp\left(-\frac{\|i - i'\|}{\sigma_{\mathrm{l}}}\right) \quad (8)$$

where σ_{EX} and σ_{l} are two hyper-parameters to determine the amplitude and length ranges of EX kernel. On the basis of EX kernel, another popular kernel named squared-exponential (SE) kernel with two hyper-parameters (σ_{SE} and σ_{l}) can be derived as:

$$k_{SE}(i, i') = \sigma_{\mathrm{SE}}^2 exp\left(-\frac{\|i - i'\|^2}{2\sigma_{\mathrm{l}}^2}\right) \quad (9)$$

It should be known that both EX and SE kernels belong to the fundamental type in that the correlation between two various observations is purely affected by $i - i'$, which would result in the smooth distributions. This may be

too strict for predicting battery electrode property with highly nonlinear behaviour. To further improve the nonlinear capture capability of GPR, two powerful covariance functions named Matern3/2 kernel and Matern5/2 kernel as expressed in Eq. (9) and (10) can be adopted.

$$k_{M3/2}(i,i') = \sigma_{\mathrm{M3/2}}^2\left(1+\frac{\sqrt{3}\|i-i'\|}{\sigma_{\mathrm{l}}}\right)exp\left(-\frac{\sqrt{3}\|i-i'\|}{\sigma_{\mathrm{l}}}\right) \tag{10}$$

$$k_{M5/2}(i,i') = \sigma_{\mathrm{M5/2}}^2\left(1+\frac{\sqrt{5}\|i-i'\|}{\sigma_{\mathrm{l}}}+\frac{5\|i-i'\|^2}{3\sigma_{\mathrm{l}}^2}\right)exp\left(-\frac{\sqrt{5}\|i-i'\|}{\sigma_{\mathrm{l}}}\right) \tag{11}$$

where $\sigma_{\mathrm{M3/2}}$ and $\sigma_{\mathrm{M5/2}}$ stand for the hyper-parameters to control the amplitudes of Matern3/2 kernel and Matern5/2 kernel, respectively. σ_{l} is another hyper-parameter.

In real applications with multiple input variables, although the EX, SE and Matern kernels have simple structure associated with a few hyper-parameters, they present limitations to fully capture the highly nonlinear relations. In our study, the inputs of interest are composed of four different IFVs (MC, STLR, CG, and viscosity). In order to well extract these input items as well as enhance the regression performance, conventional kernels including EX, SE, Matern3/2 and Matern5/2 are further reconstructed with the powerful ARD structures as ARDEX, ARDSE, ARDMatern3/2 and ARDMatern5/2 (Rasmussen and Williams, 2006), with the form of Eqs. (12), (13), (14) and (15), respectively.

$$k_{ARDEX}(i,i') = \sigma_{\mathrm{EX}}^2\, exp\left(-\sum_{c=1}^{C}\frac{\|i_c-i_c'\|}{\sigma_c}\right) \tag{12}$$

$$k_{ARDSE}(i,i') = \sigma_{\mathrm{SE}}^2\, exp\left(-\frac{1}{2}\sum_{c=1}^{C}\frac{\|i_c-i_c'\|^2}{\sigma_c^2}\right) \tag{13}$$

$$\begin{cases} k_{ARDM3/2}(i,i') = \sigma_{\mathrm{M3/2}}^2\left(1+\sqrt{3}r\right)\exp\left(-\sqrt{3}r\right) \\ r = \sqrt{\sum_{c=1}^{C}\frac{\|i_c-i_c'\|^2}{\sigma_c^2}} \end{cases} \tag{14}$$

$$\begin{cases} k_{ARDM5/2}(i,i') = \sigma_{\mathrm{M5/2}}^2\left(1+\sqrt{5}r+\frac{5}{3}r^2\right)\exp\left(-\sqrt{5}r\right) \\ r = \sqrt{\sum_{c=1}^{C}\frac{\|i_l-i_l'\|^2}{\sigma_c^2}} \end{cases} \tag{15}$$

As mentioned in related research (Jacobs, 2012; Zhang et al., 2020; Zhao et al., 2019), ARD kernel can be seen as a powerful tool to conduct features extraction and improve prediction performance in many industrial applications. By using the ARD kernel for battery electrode mass load prediction, the effects of input features (MC, STLR, viscosity and CG) on the prediction can be automatically adjusted by optimising and fixing various length scales for their hyper-parameters σ_c (here $c = 1{:}\,C$, C is the total number of input items), yielding a sparse and explanatory subset of features. It should be known that for battery production applications, the effects of feature variables on the electrode mass load could not be always same. In comparison with the conventional kernels have only one length scale for all hyper-parameters, various predictors with different length scales of hyper-parameters in ARD kernels have the potential to reduce the effects of less irrelevant feature variables and improve prediction accuracy in theory. Besides, the value of σ_c could also reflect the importance and effect of related input item. A larger σ_c results in a lower relevancy which indicates a less impact on the regression output. Benefited from this, ARD kernels could be utilised as efficient tools for direct feature variable selection. Through adopting ARD kernels for battery electrode mass load predictions, the importance weights of interested features are able to be quantified as well as selected, while the regression performance and generalization could be also enhanced by setting various length-scales of σ_c. Detailed procedures of designing ARD kernels for battery intermediate product feature selection would be described in the next subsection.

3.3 ARD kernel for feature selection

As the electrode production chain is constituted of multiple intermediate processes involving numerous IFVs and CPs, establishing a data-driven model to present effective regression predictions and quantify the importance weight of feature variables is significantly attractive.

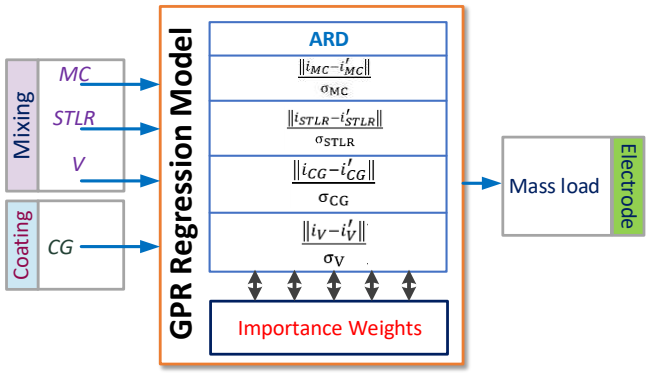


Fig. 2. GPR-based machine learning model to predict electrode mass load and analyse feature importance weights.

For the electrode production chain, the mixing and coating steps play the important roles in determining the electrode mass load, further affecting the final battery performance. In our study, to well capture the underlying correlations among feature variables and electrode mass load, a GPR-based machine learning model is designed and illustrated in Fig. 2, while the related ARD structure-based kernels would be derived as follows:

$$k_{ARDEX}(i,i') = \sigma_{\mathrm{EX}}^2 \, exp\left[-\left(\frac{\|i_{MC}-i'_{MC}\|}{\sigma_{\mathrm{MC}}} + \frac{\|i_{STLR}-i'_{STLR}\|}{\sigma_{\mathrm{STLR}}} + \frac{\|i_{CG}-i'_{CG}\|}{\sigma_{\mathrm{CG}}} + \frac{\|i_V-i'_V\|}{\sigma_{\mathrm{V}}}\right)\right] \tag{16}$$

$$k_{ARDSE}(i,i') = \sigma_{\mathrm{SE}}^2 \, exp\left[-\frac{1}{2}\left(\frac{\|i_{MC}-i'_{MC}\|^2}{\sigma_{\mathrm{MC}}^2} + \frac{\|i_{STLR}-i'_{STLR}\|^2}{\sigma_{\mathrm{STLR}}^2} + \frac{\|i_{CG}-i'_{CG}\|^2}{\sigma_{\mathrm{CG}}^2} + \frac{\|i_V-i'_V\|^2}{\sigma_{\mathrm{V}}^2}\right)\right] \tag{17}$$

$$\begin{cases} k_{ARDM3/2}(i,i') = \sigma_{\mathrm{M3/2}}^2 \left(1+\sqrt{3}r\right) exp\left(-\sqrt{3}r\right) \\ r = \sqrt{\frac{\|i_{MC}-i'_{MC}\|^2}{\sigma_{\mathrm{MC}}^2} + \frac{\|i_{STLR}-i'_{STLR}\|^2}{\sigma_{\mathrm{STLR}}^2} + \frac{\|i_{CG}-i'_{CG}\|^2}{\sigma_{\mathrm{CG}}^2} + \frac{\|i_V-i'_V\|^2}{\sigma_{\mathrm{V}}^2}} \end{cases} \tag{18}$$

$$\begin{cases} k_{ARDM5/2}(i,i') = \sigma_{\mathrm{M5/2}}^2 \left(1+\sqrt{5}r+\frac{5}{3}r^2\right) exp\left(-\sqrt{5}r\right) \\ r = \sqrt{\frac{\|i_{MC}-i'_{MC}\|^2}{\sigma_{\mathrm{MC}}^2} + \frac{\|i_{STLR}-i'_{STLR}\|^2}{\sigma_{\mathrm{STLR}}^2} + \frac{\|i_{CG}-i'_{CG}\|^2}{\sigma_{\mathrm{CG}}^2} + \frac{\|i_V-i'_V\|^2}{\sigma_{\mathrm{V}}^2}} \end{cases} \tag{19}$$

where all ARD kernels have the same set of input items as $i = \{i_{MC}, \quad i_{STLR}, \quad i_{CG}, \quad i_V\}$. σ_{MC}, σ_{STLR}, σ_{CG}, and σ_{V} represent the hyper-parameters to reflect the relevancies and importance of feature variables including MC, STLR, CG and viscosity, respectively. Based upon these defined ARD kernels, the detailed workflow to design GPR model for the feature selection and the regression prediction of electrode mass load is illustrated in Table 1.

Table 1. Detailed workflow to design the ARD kernel-based GPR for feature selection and regression of electrode mass load

1: **Procedure:** data pre-process and ARD kernel-based GPR training
a. Remove obvious outliers of original data
b. Prepare the proper inputs-output pairs for GPR training. Here the inputs are four intermediate feature variables while output is the related electrode mass load.
c. Optimise hyper-parameters θ_{hp} within ARD kernel-based GPR to minimize its $\mathrm{L}(\theta_{hp})$ through four folds cross-validation. Afterwards, obtain the well-trained GPR model $GPR_{ARD}(\theta_{hp})$.
End procedure
2: **Procedure:** feature selections and regression of electrode mass load
a. According to the θ_{hp}, calculate the feature weights by using the equation as follows: $W = exp(-\theta_{hp})$ (20)

where $\theta_{hp} = \{\sigma_{\mathrm{MC}},\ \sigma_{\mathrm{STLR}},\ \sigma_{\mathrm{CG}},\ \sigma_{\mathrm{V}}\}$ represent the well-tuned hyper-parameters of feature variable MC, STLR,CG and viscosity, respectively. $W = \{\mathrm{W_{MC}},\ \mathrm{W_{STLR}},\ \mathrm{W_{CG}},\ \mathrm{W_{V}}\}$ are their corresponding feature weights through taking the exponential of the negative hyper-parameters.

b. Quantify the feature importance by normalizing feature weights as:

$$W = W/(\mathrm{W_{MC}} + \mathrm{W_{StLr}} + \mathrm{W_{CG}} + \mathrm{W_{V}}) \quad (21)$$

c. For a new sample i' from testing dataset, use the well-trained $GPR_{ARD}(\theta_{hp})$ to predict related electrode mass load y' as:

$$y' = GPR_{ARD}(i'|\theta_{hp}) \quad (22)$$

End procedure

It should be known that taking the exponential of the negative learned θ_{hp} and normalizing W is one common solution to make the quantified feature weights become more convenient for comparison purpose (Zhang et al., 2020; Zheng and Casari, 2018). Different weighting solutions such as the inverse of θ_{hp} would lead the final quantified weight values slightly change but it would not affect the ranking of features. The feature with higher ranking (larger normalized weight value) means this feature is more important than others in determining the prediction result of electrode mass load. The quantified weight values here cannot be utilised to reflect the specific contribution proportion of using these feature variables (Dong and Liu, 2018; Zheng and Casari, 2018). In light of this, all feature weights are uniformly quantified by using Eqs. (20) and (21) in our study. Following this workflow, the weights of input items including the MC, STLR, viscosity from mixing step and the CG from coating step can be directly quantified to reflect their importance. Then the reliable feature selections can be carried out based on these quantified feature weights. Besides, the designed ARD kernel-based GPR model is capable of generating an explanatory subset of features by setting different hyper-parameters for all inputted feature variables, further helping to improve the performance and generalization ability of electrode mass load regression.

3.4 Performance metrics

To investigate the performance of the designed ARD-kernel based GPRs in electrode mass load prediction, the following four performance metrics are utilised.

1) Training time: to evaluate and compare the computation efforts of different ARD kernels, the time during GPR training process is utilised. In our study, all GPR regression models are trained through using the GPR toolbox in

Matlab 2018b. The training time with the unit second (s) for GPR model can be conveniently obtained by using *tic*, *toc* codes from the Matlab.

2) Mean absolute error (MAE): supposing N represents the total amount of regression samples, y_j are the actual testing values while $\hat{y}_j$stand for the regression values from GPR, then the MAE can be obtained to evaluate the accuracy of predictions as follows:

$$\mathrm{MAE} = \frac{1}{N}\sum_{j=1}^{N}\left|y_j - \hat{y}_j\right| \tag{23}$$

3) Root mean square error (RMSE): RMSE is another typical indicator to illustrate the deviations between the regression and actual testing values as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(y_j - \hat{y}_j\right)^2} \tag{24}$$

4. Results and discussions

To well analyse and evaluate the feature selection and regression performance of GPR with various ARD kernels, two case studies including the one with all features as well as another with reduced features are carried out in this section.

4.1 Case study 1: all features

For case study 1, to quantify the weights of interested feature variables on the regression of electrode mass load, all feature items including MC, STLR, CG and viscosity are utilised as the inputs for GPR models. The regression performance of all four ARD kernels are evaluated using the four-fold cross-validation. Fig. 3 and Table 2 illustrate the regression results and the corresponding performance metrics for the four ARD kernel-based GPRs, respectively.

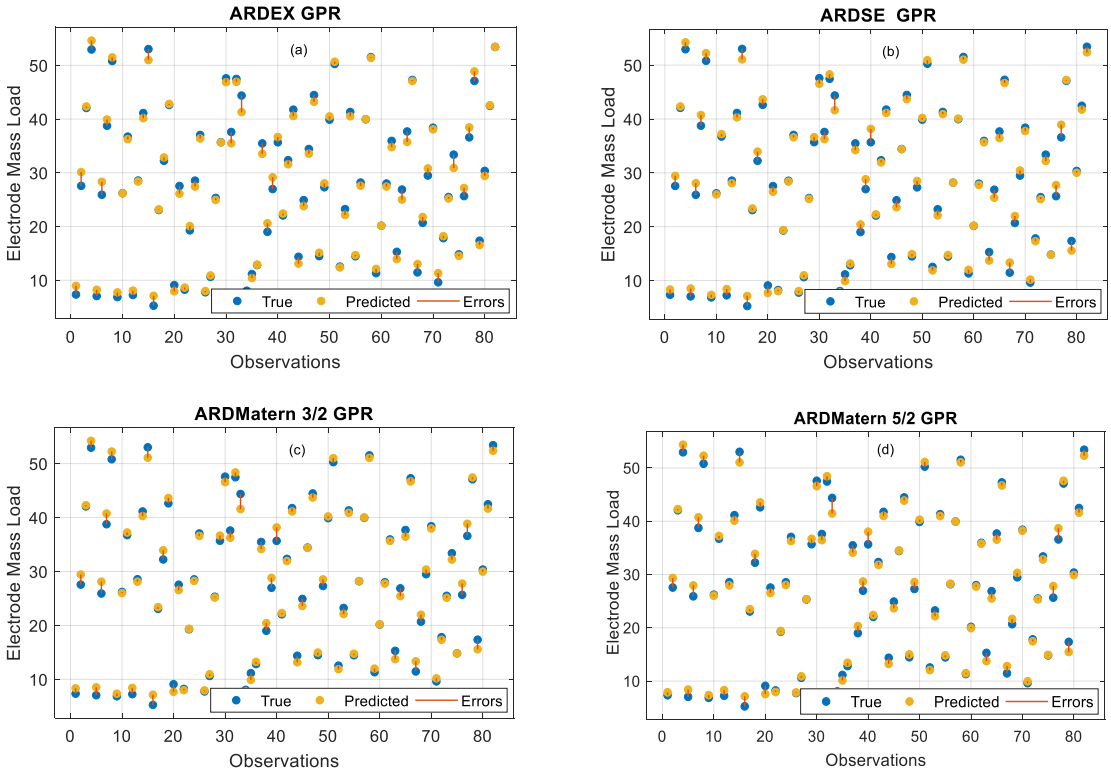


Fig. 3. Regression results through using all feature items for battery electrode mass load predictions: (a) ARDEX kernel, (b) ARDSE kernel, (c) ARDMatern3/2 kernel, (d) ARDMatern5/2 kernel.

Obviously, most observations in Fig. 3 agree well with the outputs from four GPRs, indicating that the used ARD kernels can achieve satisfactory prediction results for the electrode mass load. Quantitatively, due to the simplest structure, ARDEX kernel-based GPR can be well-trained within 13.547 s, while its accuracy is the worst with 1.177mg/cm^2 RMSE. In contrary, ARDMatern5/2 kernel-based GPR results in the longest training time of 14.487s (6.9% increase) but its RMSE value is the lowest with 1.084 mg/cm^2 (by decreasing 7.9%). However, all training time are within 14.5s, which implies that the computational efforts of four ARD kernels are acceptable. In conclusion, through using all feature variables as the input items for ARD kernel-based GPR, expected accuracy can be achieved for the prediction of battery electrode mass load.

Table 2. Performance metrics for regressions through using all feature items

Kernel types	ARDEX	ARDSE	ARDMatern3/2	ARDMatern5/2
Training time [s]	13.547	13.687	14.154	14.487
MAE [mg/cm²]	0.946	0.913	0.911	0.875
RMSE [mg/cm²]	1.177	1.126	1.124	1.084

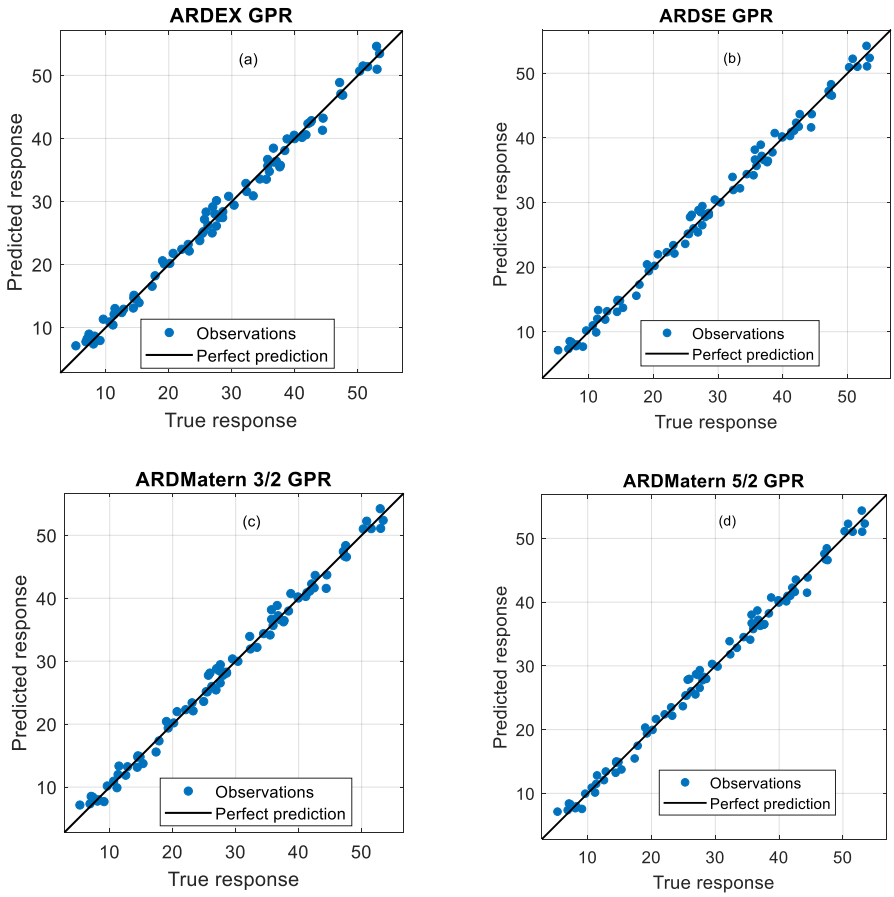


Fig. 4. Predicted versus actual plots through using all feature items: (a) ARDEX kernel, (b) ARDSE kernel, (c) ARDMatern3/2 kernel, (d) ARDMatern5/2 kernel.

To further reflect the deviations of regression results, the predicted versus actual plots (PVAPs) for all ARD kernel-based GPRs are shown in Fig. 4. It should be known that in the PVAPs, for the observations on the left or right of plots, the furthest from the average value would produce the most leverages and effectively pull the prediction line toward that observation. For a good regression model, the observations should cluster around the perfect prediction line. According to Fig. 4, all observations are clustered around the perfect prediction lines without large outliers, indicating that the ARD kernel-based GPRs are capable of achieving satisfactory regression results with a few deviations for the all feature case.

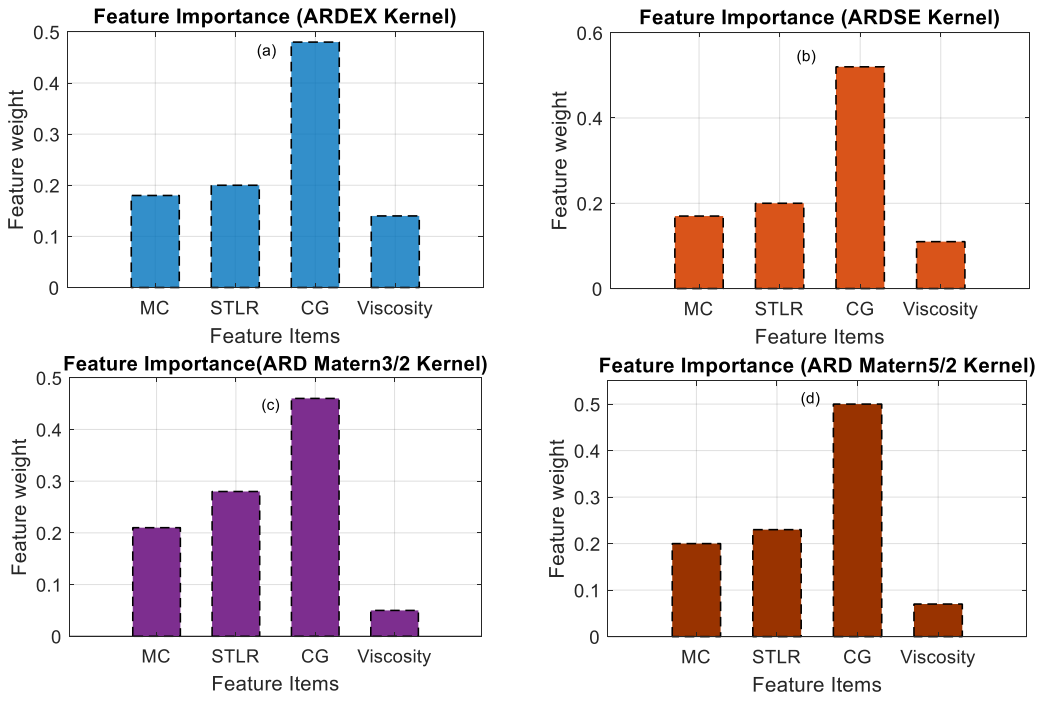


Fig. 5. Obtained feature weights through using various ARD kernels: (a) ARDEX kernel, (b) ARDSE kernel, (c) ARDMatern3/2 kernel, (d) ARDMatern5/2 kernel.

Next, according to the well-tuned GPR models, the normalized feature weights from hyper-parameters of all four ARD kernels are plotted in Fig. 5. Interestingly, although there exists difference among each feature variable, the trend of feature weights is similar for all ARD kernels. Specifically, CG provides the largest feature weight, which is nearly twice as large as the STLR. In contrary, viscosity presents the smallest feature weight (nearly five times less than the CG). The feature weights of STLR are all a slightly larger than those of MC for these ARD kernel cases. This finding signifies that among these four feature variables, CG is the most important one and must be selected for electrode mass load prediction. STLR and MC are the second and third important feature items. In contrast, viscosity makes the least contribution to predicting the battery electrode mass load.

4.2 Case study 2: reduced features

According to the above case study, the weights of all these four feature items are quantified from the hyper-parameters within ARD kernel. Then the importance of these feature variables can be successfully ranked. In accordance with the analysis in Section 4.1, reduced feature tests accounting for less features are carried out in this section to further evaluate the effectiveness of feature ranking. Specifically, the most important three feature variables including CG, STLR and MC are utilised as the input items to evaluate the performance of all ARD

kernel-based GPRs again with four-fold cross-validation. As the number of feature items in our study is limited as four, removing the least important feature item is enough to evaluate the sensitivity analysis of feature variables in such cases (Zheng and Casari, 2018). When more data of feature variables from battery production chain are collected to generate a larger feature pool, a solution by setting the threshold for eliminating the particular feature variables with lower quantified weights can be utilised due to its flexibility. The regression results through using these reduced feature items and the corresponding performance indicators are presented in Fig. 6 and Table 3, respectively.

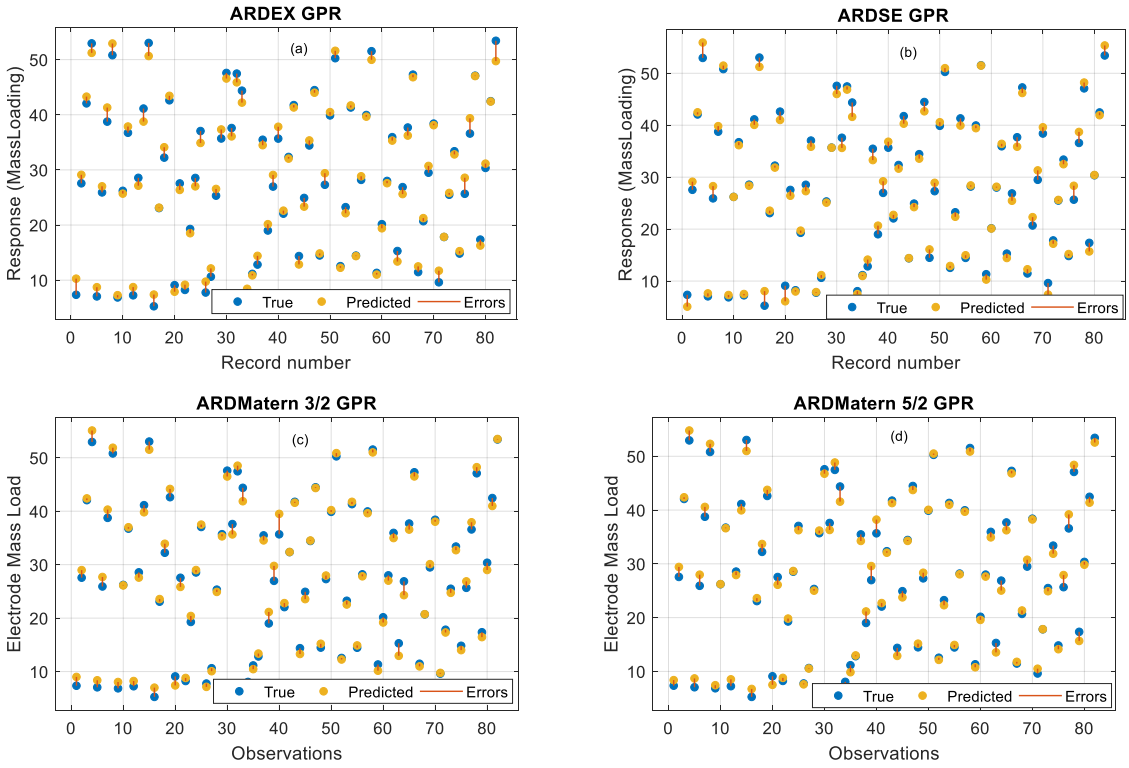


Fig. 6. Regression results through using reduced feature items for battery electrode mass load predictions: (a) ARDEX kernel, (b) ARDSE kernel, (c) ARDMatern3/2 kernel, (d) ARDMatern5/2 kernel.

Table 3. Performance metrics for regressions through using reduced feature items

Kernel types	ARDEX	ARDSE	ARDMatern3/2	ARDMatern5/2
Training time [s]	11.237	11.632	12.986	13.124
MAE [mg/cm²]	1.197	1.066	0.997	0.985
RMSE [mg/cm²]	1.433	1.324	1.222	1.204

It can be noticed that removing one feature item (viscosity) contributes to reducing the training time of all ARD kernel cases. The maximum training time is 13.124 s for ARDMatern5/2 kernel, which is 9.4% less than its counterpart with full feature. As observed from Fig. 6, although there exist a few more obvious mismatched observations, most observations can be still well predicted by the ARD kernel-based GPRs. Quantitatively, the worst regression performance is observed from ARDEX, with 1.433mg/cm^2 RMSE (17.7% increase in comparison with that of all feature case). The ARDMatern kernels still achieve the best regression performance, which are 1.222mg/cm^2 RMSE for ARDMatern3/2 (8.7% increase) and 1.204mg/cm2 RMSE (9.9% increase) for ARDMatern5/2, respectively. These results imply that the GPR with ARDMatern kernels provide better accuracy for electrode mass load prediction in both all feature case and reduced feature case.

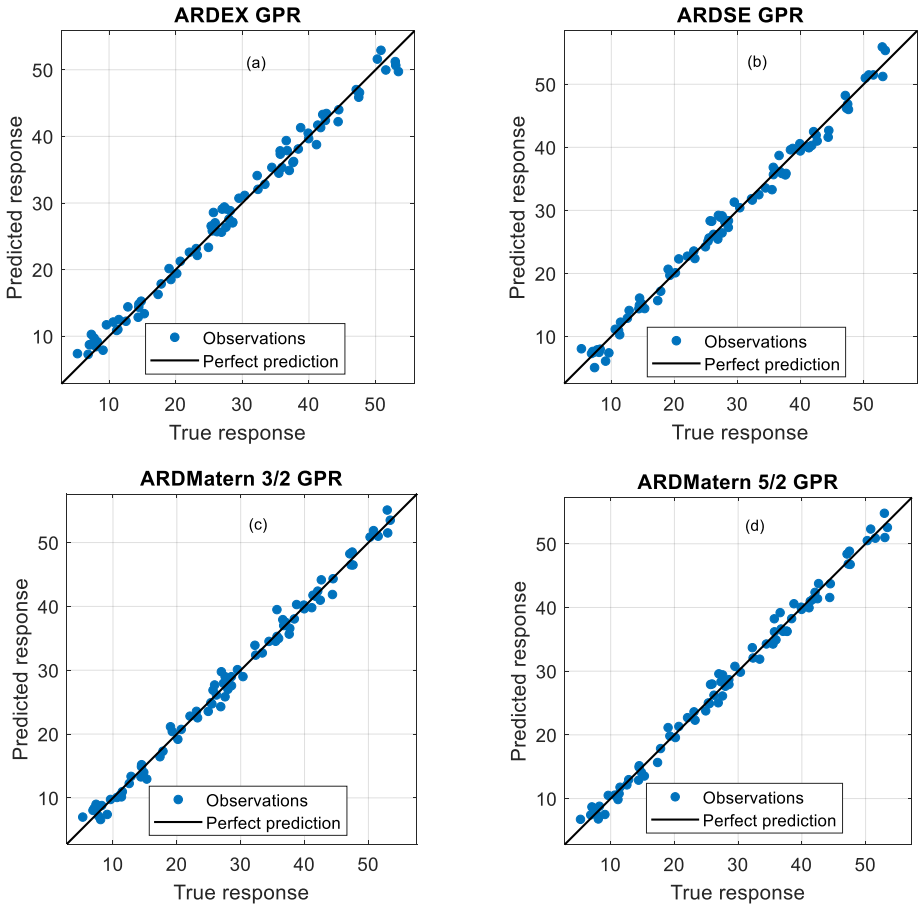


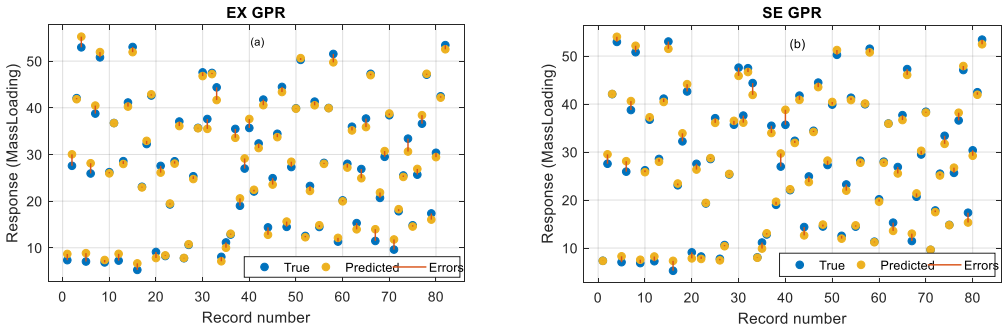
Fig. 7. Predicted versus actual plots through using reduced feature items: (a) ARDEX kernel, (b) ARDSE kernel, (c) ARDMatern3/2 kernel, (d) ARDMatern5/2 kernel.

To reflect the deviations of these regression results, the PVAPs using GPRs with reduced feature variables are shown in Fig. 7. In comparison with the PVAPs using all features, the deviations of observations in the reduced

feature case spread over a wide uncertain region. However, all observations are still dispersed around the perfect prediction lines closely, which means that none obvious prediction outliers are generated in all ARD kernel cases. In summary, after removing the viscosity item with the lowest feature weight, the ARD-kernel based GPRs can still provide satisfactory regression results for electrode mass load. These results imply that the feature importance weights can be successfully quantified by the designed ARD kernel. For the battery production chain that contains numerous feature variables, the proposed ARD kernel-based GPR is able to not only reduce the feature dimension but also guarantee the performance of electrode property regression, further helping to promote the efficiencies.

In our study, following the same flow in many feature engineering applications (Dong and Liu, 2018; Zheng and Casari, 2018), the ARD-kernel based GPR with all features is first utilised to get importance ranking of these features and then the reduced features GPR is adopted for further evaluating the correctness of selected features. Due to the number of features is four, we only remove a feature variable with lower importance weight in the reduced feature GPR case, further resulting the training time is not highly reduced. As battery production chain is a highly-complicated process with numerous feature variables in total, monitoring all feature variables and using full features to build a model is time-consuming and unrealistic. In light of this, when more feature data within battery production chain are available, the advantages of using a reduced features GPR would become more obvious as it could help to remove more irrelevant feature items and reduce the training time without sacrificing much accuracy.

4.3 Performance comparison with conventional kernels



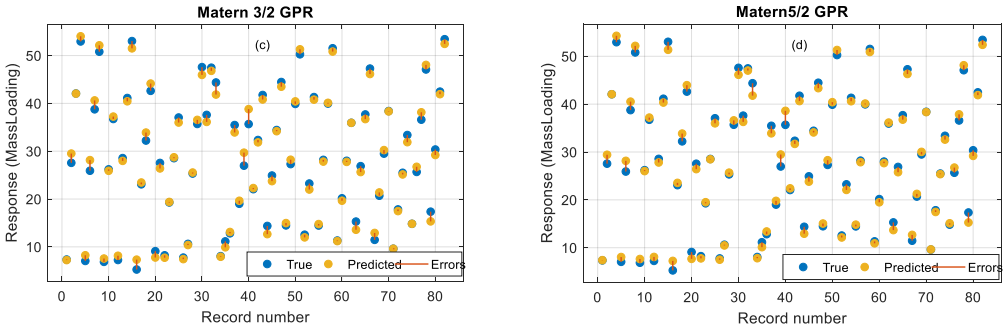


Fig. 8. Comparison results through using conventional kernels with all feature items: (a) EX kernel, (b) SE kernel, (c) Matern3/2 kernel, (d) Matern5/2 kernel.

Table 4. Performance metrics for regressions through using all feature items

Kernel types	EX	SE	Matern3/2	Matern5/2
Training time [s]	13.323	13.586	13.964	13.959
MAE [mg/cm²]	0.994	0.983	0.958	0.907
RMSE [mg/cm²]	1.317	1.292	1.184	1.151

To further reflect the prediction performance of ARD kernels, four GPR models with the related conventional kernels of EX, SE, Matern3/2 and Matern5/2 are adopted as the benchmarks for comparison purpose. According to the specific forms as illustrated in Eqs. (8), (9), (10) and (11), these convention kernels without ARD structure cannot directly quantify the feature importance as the hyper-parameters of all four feature items within the kernel are same. Fig. 8 and Table 4 illustrate the regression results and corresponding performance metrics of using all four feature items as inputs to the conventional kernel-based GPRs after four-folds cross-validation. It can be seen that EX kernel shows the worst results, while Matern3/2 and Matern5/2 provide better regression results. Here the RMSEs of all conventional kernels are a slightly worse than the related ARD kernels (5% increase). Therefore, apart from the superiority in terms of interpretability of feature importance, ARD kernel-based GPRs also present competent performance in the prediction of battery electrode mass load.

4.4 Implications for theory and practice on battery production

The ARD kernels-based GPR machine learning framework proposed in this study has extensive application prospects for the predictions of battery product properties and the sensitivity analysis of various feature variables in battery production application. For instance, the properties of manufactured electrode play a key role in

determining the battery final performance (energy or power densities, maximal capacity and service life), which must be carefully predicted and managed. Besides, the battery production chain contains many intermediate stages such as mixing, coating, drying and calendaring. All these stages cover multiple disciplines such as mechanical, chemical and electrical engineering, making the traditional sensitivity analysis by trial and error way is significantly time-consuming. Another complication is that battery production chain has numerous strong-coupled feature variables in total, monitoring all feature variables and using full features to build a model is inefficient and unrealistic. In light of this, this paper proposes the idea to simultaneously predict battery electrode property and analyse feature importance through machine learning approach. When more feature data within battery production chain are available, the proposed framework could help to remove more irrelevant feature items and reduce the training time without sacrificing much accuracy. The battery manufacturer can utilise the feature importance information and property prediction results from our model to redesign the monitor variables and readjust the process parameters to improve the manufacturing performance, so as to ensure the effectiveness of monitored feature variables and the satisfactory qualities of cleaner battery production.

5. Conclusion

As the electrode production plays a key role in affecting the results and performance for cleaner battery production, the efficient electrode mass load prediction and feature selection of battery intermediate products are presented in this study. An advanced machine learning approach, based on the GPR technology with various ARD kernels, is developed to effectively predict the battery electrode mass load. Meanwhile, the importance weights of four feature variables of interest in the mixing and coating steps are also directly quantified through analysing the hyper-parameters within ARD kernels. According to the quantitative comparison and sensitivity analysis of results, the main conclusions can be summarized as:

1) The ARD kernel-based GPRs are capable of not only predicting electrode mass load (here the RMSEs are at least 5% better than those from conventional kernel-based GPRs) but also well quantifying the importance weights for all four feature variables (MC, STLR, CG and viscosity). The obtained importance ranking can well benefit the feature selection and model dimension reduction for efficient battery production.

2) Among four ARD kernel functions, the ARDMatern5/2 kernel achieves the best regression performance for the cases of either using all feature items (at least 4% improved) or reduced feature items as inputs (at least 2% improved), making it more suitable for property prognostics of battery electrode production.

3) CG from coating stage is most crucial for electrode mass load regression among four feature variables of interest. Here the quantified importance value of CG is twice more than others. Therefore, CG should be well monitored and controlled to obtain electrode product with reliable performance.

4) On the contrary, viscosity from mixing stage makes the least contribution to the electrode mass load regression (9.9% RMSE increases). To further improve the performance of electrode mass load prognostics in battery production, the battery manufacturer should pay more attention to MC and STLR when operating the mixing stage.

This is the first known application by deriving ARD kernel-based GPR framework to directly and simultaneously predict battery electrode property as well as analyse feature importance. As the designed machine learning framework in our study belongs to a data-driven approach with the superiority in terms of accuracy and flexibility, it can be conveniently extended to analyse other variables within the battery production chain when the corresponding data are available. Our future work would focus on designing extra experiments to generate more available data from other battery production process (i.e., mixing, drying and calendaring), and then to conduct reliable sensitivity analysis and feature selection for more variables within battery production chain, further benefitting the cleaner battery production and optimal energy storage system design.

Acknowledgements

This work is supported by the EPSRC under grant EP/R030243/1, the National Key Research and Development Project under Grant 2018YFB 1700503, and the Natural Science Foundation of Guangdong under Grant No. 2018A 030310671. The authors would like to thank professor Jihong Wang from School of Engineering, University of Warwick for fruitful discussions and proofreading.

Reference

Chouchane, M., Rucci, A., Lombardo, T., Ngandjong, A.C., Franco, A.A., 2019. Lithium ion battery electrodes predicted from manufacturing simulations: Assessing the impact of the carbon-binder spatial location on the electrochemical performance. JPS 444, 227285. https://doi.org/10.1016/j.jpowsour.2019.227285

Cunha, R.P., Lombardo, T., Primo, E.N., Franco, A.A., 2020. Artificial Intelligence Investigation of NMC Cathode Manufacturing Parameters Interdependencies. Batteries & Supercaps 3(1), 60-67. https://doi.org/10.1002/batt.201900135

Dahodwalla, H., Herat, S., 2000. Cleaner production options for lead-acid battery manufacturing industry. J. Clean. Prod. 8(2), 133-142. https://doi.org/10.1016/S0959-6526(99)00314-5

Dong, G., Liu, H., 2018. Feature Engineering for Machine Learning and Data Analytics. CRC Press.

Günther, T., Schreiner, D., Metkar, A., Meyer, C., Kwade, A., Reinhart, G., 2020. Classification of Calendering - Induced Electrode Defects and Their Influence on Subsequent Processes of Lithium - Ion Battery Production. Energy Technol. 8(2), 1900026. https://doi.org/10.1002/ente.201900026

He, H., Tian, S., Tarroja, B., Ogunseitan, O.A., Samuelsen, S., Schoenung, J.M., 2020. Flow battery production: Materials selection and environmental impact. J. Clean. Prod. 269, 121740. https://doi.org/10.1016/j.jclepro.2020.121740

Hoffmann, L., Grathwol, J.-K., Haselrieder, W., Leithoff, R., Jansen, T., Dilger, K., Dröder, K., Kwade, A., Kurrat, M., 2019. Capacity Distribution of Large Lithium-Ion Battery Pouch Cells in Context with Pilot Production Processes. Energy Technol. 8. https://doi.org/10.1002/ente.201900196

Jacobs, J.P., 2012. Bayesian Support Vector Regression With Automatic Relevance Determination Kernel for Modeling of Antenna Input Characteristics. IEEE Trans. Antennas Propag. 60(4), 2114-2118. https://doi.org/10.1109/TAP.2012.2186252

Knoche, T., Surek, F., Reinhart, G., 2016. A Process Model for the Electrolyte Filling of Lithium-ion Batteries. Procedia CIRP 41, 405-410. https://doi.org/10.1016/j.procir.2015.12.044

Kwade, A., Haselrieder, W., Leithoff, R., Modlinger, A., Dietrich, F., Droeder, K., 2018. Current status and challenges for automotive battery production technologies. Nat. Energy 3(4), 290-300. https://doi.org/10.1038/s41560-018-0130-3

Lee, S.S., Kim, T.H., Hu, S.J., Cai, W.W., Abell, J.A., 2010. Joining Technologies for Automotive Lithium-Ion Battery Manufacturing: A Review. ASME 2010 International Manufacturing Science and Engineering Conference, Volume 1, 541-549. https://doi.org/10.1115/MSEC2010-34168

Li, Y., Liu, K., Foley, A.M., Zülke, A., Berecibar, M., Nanini-Maury, E., Van Mierlo, J., Hoster, H.E., 2019. Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. Renew. Sust. Energ. Rev. 113, 109254. https://doi.org/10.1016/j.rser.2019.109254

Lipu, M.S.H., Hannan, M.A., Hussain, A., Hoque, M.M., Ker, P.J., Saad, M.H.M., Ayob, A., 2018. A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations. J. Clean. Prod. 205, 115-133. https://doi.org/10.1016/j.jclepro.2018.09.065

Liu, C., Lin, J., Cao, H., Zhang, Y., Sun, Z., 2019. Recycling of spent lithium-ion batteries in view of lithium recovery: A critical review. J. Clean. Prod. 228, 801-813. https://doi.org/10.1016/j.jclepro.2019.04.304

Liu, K., Hu, X., Wei, Z., Li, Y., Jiang, Y., 2019. Modified Gaussian Process Regression Models for Cyclic Capacity Prediction of Lithium-Ion Batteries. IEEE Trans. Transp. Electrification 5(4), 1225-1236. https://doi.org/10.1109/TTE.2019.2944802

Liu, K., Li, Y., Hu, X., Lucu, M., Widanalage, D., 2019. Gaussian Process Regression with Automatic Relevance Determination Kernel for Calendar Aging Prediction of Lithium-ion Batteries. IEEE Transactions on Industrial Informatics PP, 1-1. https://doi.org/10.1109/TII.2019.2941747

Liu, K., Shang, Y., Ouyang, Q., Widanage, W.D., 2020. A Data-driven Approach with Uncertainty Quantification for Predicting Future Capacities and Remaining Useful Life of Lithium-ion Battery. ITIE, 1-1. https://doi.org/10.1109/TIE.2020.2973876

Liu, K., Zou, C., Li, K., Wik, T., 2018. Charging Pattern Optimization for Lithium-Ion Batteries With an Electrothermal-Aging Model. IEEE Transactions on Industrial Informatics 14(12), 5463-5474. https://doi.org/10.1109/TII.2018.2866493

Maia, L.K.K., Drünert, L., La Mantia, F., Zondervan, E., 2019. Expanding the lifetime of Li-ion batteries through optimization of charging profiles. J. Clean. Prod. 225, 928-938. https://doi.org/10.1016/j.jclepro.2019.04.031

Meng, J., Stroe, D., Ricco, M., Luo, G., Swierczynski, M., Teodorescu, R., 2019. A Novel Multiple Correction Approach for Fast Open Circuit Voltage Prediction of Lithium-Ion Battery. ITEnC 34(2), 1115-1123. https://doi.org/10.1109/TEC.2018.2880561

Ouyang, Q., Wang, Z., Liu, K., Xu, G., Li, Y., 2020. Optimal Charging Control for Lithium-Ion Battery Packs: A Distributed Average Tracking Approach. IEEE Transactions on Industrial Informatics 16(5), 3430-3438. https://doi.org/10.1109/TII.2019.2951060

Pan, Y., Feng, X., Zhang, M., Han, X., Lu, L., Ouyang, M., 2020. Internal short circuit detection for lithium-ion battery pack with parallel-series hybrid connections. J. Clean. Prod. 255, 120277. https://doi.org/10.1016/j.jclepro.2020.120277

Pang, Y., Cao, Y., Chu, Y., Liu, M., Snyder, K., MacKenzie, D., Cao, C., 2019. Additive Manufacturing of Batteries. Adv. Funct. Mater., 1906244. https://doi.org/10.1002/adfm.201906244

Rasmussen, C.E., Williams, C.K.I., 2006. Gaussian Processes for Machine Learning. University Press Group Limited.

Richardson, R.R., Birkl, C.R., Osborne, M.A., Howey, D.A., 2019. Gaussian Process Regression for In Situ Capacity Estimation of Lithium-Ion Batteries. IEEE Transactions on Industrial Informatics 15(1), 127-138. https://doi.org/10.1109/TII.2018.2794997

Saw, L.H., Ye, Y., Tay, A.A.O., 2016. Integration issues of lithium-ion battery into electric vehicles battery pack. J. Clean. Prod. 113, 1032-1045. https://doi.org/10.1016/j.jclepro.2015.11.011

Schnell, J., Nentwich, C., Endres, F., Kollenda, A., Distel, F., Knoche, T., Reinhart, G., 2019. Data mining in lithium-ion battery cell production. JPS 413, 360-366. https://doi.org/10.1016/j.jpowsour.2018.12.062

Schnell, J., Reinhart, G., 2016. Quality Management for Battery Production: A Quality Gate Concept. Procedia CIRP 57, 568-573. https://doi.org/10.1016/j.procir.2016.11.098

Schünemann, J.-H., Dreger, H., Bockholt, H., Kwade, A., 2016. Smart Electrode Processing for Battery Cost Reduction. ECS Transactions 73, 153-159. https://doi.org/10.1149/07301.0153ecst

Shang, Y., Liu, K., Cui, N., Wang, N., Li, K., Zhang, C., 2019. A Compact Resonant Switched-Capacitor Heater for Lithium-Ion Battery Self-Heating at Low Temperatures. ITPE PP, 1-1. https://doi.org/10.1109/TPEL.2019.2954703

Tagade, P., Hariharan, K.S., Ramachandran, S., Khandelwal, A., Naha, A., Kolake, S.M., Han, S.H., 2020. Deep Gaussian process regression for lithium-ion battery health prognosis and degradation mode diagnosis. JPS 445, 227281. https://doi.org/10.1016/j.jpowsour.2019.227281

Tang, X., Liu, K., Wang, X., Liu, B., Gao, F., Widanage, W.D., 2019. Real-time aging trajectory prediction using a base model-oriented gradient-correction particle filter for Lithium-ion batteries. JPS 440, 227118. https://doi.org/10.1016/j.jpowsour.2019.227118

Thiede, S., Turetskyy, A., Kwade, A., Kara, S., Herrmann, C., 2019. Data mining in battery production chains towards multi-criterial quality prediction. CIRP Annals 68. https://doi.org/10.1016/j.cirp.2019.04.066

Tian, H., Qin, P., Li, K., Zhao, Z., 2020. A review of the state of health for lithium-ion batteries: Research status and suggestions. J. Clean. Prod. 261, 120813. https://doi.org/10.1016/j.jclepro.2020.120813

Turetskyy, A., Thiede, S., Thomitzek, M., von Drachenfels, N., Pape, T., Herrmann, C., 2019. Toward Data - Driven Applications in Lithium - Ion Battery Cell Manufacturing. Energy Technol. https://doi.org/10.1002/ente.201900136

Wanner, J., Weeber, M., Birke, K.P., Sauer, A., 2019. Quality Modelling in Battery Cell Manufacturing Using Soft Sensoring and Sensor Fusion - A Review. 2019 9th International Electric Drives Production Conference (EDPC), 1-9. https://doi.org/10.1109/EDPC48408.2019.9011847

Wei, Z., Zhao, J., Xiong, R., Dong, G., Pou, J., Tseng, K.J., 2019. Online Estimation of Power Capacity With Noise Effect Attenuation for Lithium-Ion Battery. ITIE 66(7), 5724-5735. https://doi.org/10.1109/TIE.2018.2878122

Xiong, R., Chen, H., Wang, C., Sun, F., 2018. Towards a smarter hybrid energy storage system based on battery and ultracapacitor - A critical review on topology and energy management. J. Clean. Prod. 202, 1228-1240. https://doi.org/10.1016/j.jclepro.2018.08.134

Yang, D., Zhang, X., Pan, R., Wang, Y., Chen, Z., 2018. A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve. JPS 384, 387-395. https://doi.org/10.1016/j.jpowsour.2018.03.015

Yang, R., Xiong, R., He, H., Chen, Z., 2018. A fractional-order model-based battery external short circuit fault diagnosis approach for all-climate electric vehicles application. J. Clean. Prod. 187, 950-959. https://doi.org/10.1016/j.jclepro.2018.03.259

Zhang, Y., Tang, Q., Zhang, Y., Wang, J., Stimming, U., Lee, A.A., 2020. Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning. Nat. Commun. 11(1), 1706. https://doi.org/10.1038/s41467-020-15235-7

Zhang, Y., Xiong, R., He, H., Pecht, M., 2019. Validation and verification of a hybrid method for remaining useful life prediction of lithium-ion batteries. J. Clean. Prod. 212, 240-249. https://doi.org/10.1016/j.jclepro.2018.12.041

Zhao, J., Chen, L., Pedrycz, W., Wang, W., 2019. Variational Inference-Based Automatic Relevance Determination Kernel for Embedded Feature Selection of Noisy Industrial Data. ITIE 66(1), 416-428. https://doi.org/10.1109/TIE.2018.2815997

Zheng, A., Casari, A., 2018. Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists. O'Reilly Media.

Zhou, Z., Duan, B., Kang, Y., Shang, Y., Cui, N., Chang, L., Zhang, C., 2020. An efficient screening method for retired lithium-ion batteries based on support vector machine. J. Clean. Prod. 267, 121882. https://doi.org/10.1016/j.jclepro.2020.121882

Zhou, Z., Kang, Y., Shang, Y., Cui, N., Zhang, C., Duan, B., 2019. Peak power prediction for series-connected LiNCM battery pack based on representative cells. J. Clean. Prod. 230, 1061-1073. https://doi.org/10.1016/j.jclepro.2019.05.144