Journal of Energy Storage 84 (2024) 110783

Contents lists available at ScienceDirect

Journal of Energy Storage

journal homepage: www.elsevier.com/locate/est

Research papers

Unveiling the performance impact of module level features on parallel-connected lithium-ion cells via explainable machine learning techniques on a full factorial design of experiments

Gabriele Piombo [a,*,1], Simone Fasolato [b,1], Robert Heymer [a], Marc Hidalgo [a], Mona Faraji Niri [a], Simona Onori [c], James Marco [a]

[a] Warwick Manufacturing Group, University of Warwick, Coventry, United Kingdom
[b] Department of Electrical, Computer and Biomedical Engineering, University of Pavia, Pavia, Italy
[c] Department of Energy Science and Engineering, Stanford University, Stanford, USA

ARTICLE INFO

Keywords:
Lithium-ion battery
Parallel-connected cells
Cell-to-cell parameters variation
Current and temperature imbalance
Explainable machine learning

ABSTRACT

Parallel string performance imbalances are unavoidable due to manufacturing-related cell-to-cell inhomogeneities (e.g. capacity, internal resistance), suboptimal pack and cooling system design. Understanding the most important features at a single cell and module level influencing the heterogeneity propagation inside the modules/packs is therefore crucial to limiting the phenomenon. In this article, a methodology combining well-established non-invasive single-cell characterisation tests with data-driven modelling tools is proposed. Two batches of twenty new, commercial NMC and NCA cells are first characterised to identify out-of-manufacture internal resistance and capacity distributions. Then, a 54 test condition full-factorial Design of Experiment campaign on four cells ladder-parallel connected modules is performed. The experiments inform about how the cells' current, State of Charge, temperature distributions and time to self-balance under 0.75C constant current discharge loads are affected by interconnection resistance, operating temperature, different chemistry combination and ageing. The multivariate linear model analysis confirms that combining NMC and NCA cells in parallel is possible both for first and second life applications. Nevertheless, mixing different chemistries and including an aged cell show a detrimental effect on the balanced performance of the module. The application of Explainable Machine Learning techniques such as SHAP, Partial Dependence Plots and Individual Conditional Expectation closes the gap between data-driven models' interpretability against traditional black box models while maintaining the advantage of capturing highly non-linear control-response relationships. According to the results, the interconnection resistance is the most relevant contributor to heterogeneous performance within the string. In the first and middle phases of the discharge, the distributions of internal resistance and capacity impact the load imbalance across the cells, respectively. Increasing the operating temperature contributes to exacerbate the thermal gradient in the string.

1. Introduction

Lithium-ion batteries are an essential technology for meeting the decarbonisation objectives in the transportation and energy sectors [1]. Depending on the application, individual cells are combined using various series and parallel architectures to form modules and packs to meet the target power and energy requirements [2]. The serial connection assures greater voltages to contain losses [3]. The parallel connection allows the package to store an appropriate quantity of energy [4].

Prior research has effectively identified and synthesised the key factors that exert the most significant impact on the efficacy of parallel cell modules [5]. The lithium-ion cells manufacturing method is constantly evolving and refining, currently providing high quality uniformity among fresh cells ensuring the adaptability to mass scale demand [6]. Despite this, variations in the characteristics of single cells are possible [7]. Manufacturing-related cell-to-cell variations could be in the form of internal resistance [8–10], capacity [11, 12], their combination [13–15] and open circuit voltage Open Circuit

* Corresponding author.
E-mail address: gabriele.piombo@warwick.ac.uk (G. Piombo).
1 Visiting Scholar at the Department of Energy Science and Engineering, Stanford University.

https://doi.org/10.1016/j.est.2024.110783
Received 15 August 2023; Received in revised form 19 December 2023; Accepted 30 January 2024
Available online 13 February 2024

Nomenclature

AI	Artificial Intelligence
AICc	Akaike Information Criterion
ALE	Accumulated Local Effects
ANOVA	Analysis of Variance
CAN	Controller Area Network
CC-CV	Constant Current Constant Voltage
CV	Cross Validation
DoE	Design of Experiments
DT	Decision Tree
GS	Grid Search
HPPC	Hybrid Pulse Power Characterisation
ICE	Individual Conditional Expectation
KPI	Key Performance Indicator
MDI	Mean Decrease in Impurity
ML	Machine Learning
MLP	Multi-Layer Perceptron
NCA	Nichel Cobalt Aluminium
NMC	Nichel Manganese Cobalt
OCV	Open Circuit Voltage
PDP	Partial Dependence Plot
RF	Random Forest
RMSE	Root Mean Squared Error
RS	Random Search
SHAP	SHapley Additive exPlanations
SiC	Silicon-doped graphite
SoC	State of Charge
SoH	State of Health
TTSB	Time To Self Balance
VIF	Variance Inflation Factor
XML	Explainable Machine Learning
XAI	Explainable Artificial Intelligence

Voltage (OCV) [16,17]. They all contribute to different forms of imbalance in parallel strings. Not only single-cell level features but also module-level characteristics including interconnection resistance [18–20], the number of cells in parallel [21,22], topology selection [23,24] and chemistry combination [25,26] have a non-negligible impact on pack performance. Operating temperature [27–29] and poor-cooling design induced thermal gradients [30,31] can also affect the uniformity of packs performance. Furthermore, improper dimensioning of electrical connections between cells might result in increased local resistance [32]. These local attributes cause load peaks that are detrimental to the system's operation as they change the operating loads on each cell [33–35]. To fully investigate the behaviour of battery packs and understand their health evolution, the view must be elevated from single component to system level.

The principal challenges that require additional exploration are parallel connection generated load current [11,36,37] and temperature [19,20] imbalance, over-time internal resistance, capacity [30,38,39] and ageing rate [5,40,41] cell-to-cell fluctuation. In particular, the phenomenon of performance imbalance leading to non-uniform ageing of individual cells has been reported in literature. Specifically, it has been noted that the prolongation of imbalances is a contributing factor to this phenomenon. The existing literature presents divergent views on this matter. While some researchers [13,14,42,43] affirmed that there exists a convergence and self-balancing attitude among parallel connected cells over time, others' [15,44–47] findings oppose this theory. So far, most of the research focus has been on individual cells' behaviour, with some experimental assessment of module connections resumed in [5,18]. Coherently, the issue of parallel cell connection leaves gaps in the knowledge and necessitates further investigation. Table 1 offers a comprehensive summary of the experimental investigations carried out on cells that are connected in parallel, extending the information available in [5,18]. The literature review enables the inclusion of pertinent details regarding the cells and test characteristics, control variables and responses. The integration of experiments and modelling has emerged as a potentially viable substitute for exclusive experimental endeavours over the last years [48]. Nevertheless, precise empirical data serves as the foundation for all modelling initiatives. The investigations are of an experimental nature, with Rumpf [49] and Hosseinzadeh [50] studies representing exceptions. Both these studies suggest analysing module behaviour differently, validating equivalent or electrochemical models at the cell level and expanding them to represent modules. The method can incorporate most control variables, including cell attributes and string properties. Thus, architecture, interconnection resistance, thermal gradients, and a number of parallel cells can be combined. Despite their cell-level precision, this methodology's main limitation is validating the models at the module level, due to the large number of factors involved. After cells reach their End-of-Life (EoL), typically taken as 80% of the original nominal capacity or 200% of the internal resistance, cells are commonly retired or recycled [51]. The environmental sustainability of battery applications strongly relies on the usable life extension, generally indicated as second life. Despite recent studies developed increasingly innovative algorithms for the selection and redirection of individual cells and entire modules for their second life [52,53], few have considered the incorporation and combination of different chemistries. As indicated in Table 1, Chang [25] conducted an experimental study that included NMC, NCA, and LMO cells to investigate their impact on heterogeneities when connected in parallel and, consequently, their suitability for second life purposes. The authors conclude that NMC and NCA cells are compatible, whereas LMO and NCA cells are not. The literature-available experimental approaches largely rely on testing small batches (up to 8) of new "18650" and pouch format cells. Few exceptions included "21700" [25] and "26650" [22,49,54] formats. Some studies aimed at increasing the generalisability of the experiments by enlarging the available statistical pool [26,39,55] or varying the connection topology [56–58]. Increasing the number of tested cells can help to estimate the distribution of the characteristics of out-of-manufacture batches and hence generalising their statistical influence on modules' performance. However, little has been done in parallel cell studies to include representative cells from new batches of cells.

Uncontrolled test bench connections and laboratory testing conditions can influence the results [18]. Interconnection resistance has been found to be a significant factor to the imbalanced performance of parallel strings and should not be conflated with contact resistance. The contact resistance is modelled as a series component with the cells and its presence can mitigate current imbalance due to cell-to-cell discrepancies in the string [58]. The interconnection resistance can instead be modelled between parallel branches and has been observed to have a detrimental effect on the overall energy balance of a string [62]. Nevertheless, few studies reported measurements of the setup contact and interconnection resistances [18,19,54,58,59]. Jocher [18] reported an analysis of the consistency of literature-employed experimental settings and types of branch current sensors and proposed a novel approach able to isolate the influence of interconnection resistance via virtual connections. Virtual connections mimic physical connections without affecting string performance, acknowledging the importance the setup can have on the load distribution. Experiments show that this method can accurately assess string performance and the extent of interconnection resistance contribution. Previous studies have predominantly utilised shunts or Hall-Effect sensors. The former affects current distribution by changing branch resistance [59], whereas the latter is less precise [13]. Fill's study [19] examined branch current measurement using a contact cable with varying resistance and acknowledged the

Table 1
Summary of the recent literature on experimental studies on parallel-connected modules.

Reference			Study characteristics								Control variables											Response variables					
Author	Ref.	Year	Cell Format	Cell Chem.	Exp. SC	Exp. Mod.	Arch.	Cell Batch	Curr. Sens.	Meas R_{Int} [mΩ]	ΔR_0	R_{Int}	$T_{Grad.}$	T	ΔC_n	OCV	$Chem.$	$No.P-cells$	$Arch$	$Load$	SOC	σI	$T_{Grad.}$	ΔSoC	$E_{Mod.}$	ΔSOH	$Ageingrate$
Al-Amin	[59]	2021	18 650	NMC	x	x	4p	4	Hall	0.06–0.083	x		x		x					x	x	x	x			x	
Baumann	[5]	2018	18 650				2p																				
Chang	[25]	2022	18 650, 21 700	NMC, NCA, LMO	x	x	2p	7	Hall	N.A.	x				x	x	x					x		x			
Chang	[17]	2020	18 650	NMC, NCA, LMO	x	x	2p, 3p	8	Hall	N.A.	x				x	x				x		x		x			
Cui	[55]	2022	18 650	NMC	x	x	2p, 3p, 4p	54	Hall	N.A.	x				x	x		x			x	x		x			
Diao	[22]	2019	18 650, 26 650	LFP, NCA	x	x	3p	6	N.A.	N.A.	x				x			x		x		x					
Fill	[19]	2019	Pouch	LCO (LNCO)	x	x	2p	2	Cable 0.3, 1.6, 4.35 [mΩ]	0.08		x		x								x	x				
Fill	[21]	2019	N.A.	N.A.	x	x	2p	2	Cable	N.A.	x				x	x		x				x					
Fill	[32]	2019	Pouch	N.A.			3s2p	6	Not meas.	N.A.																	
Fill	[60]	2020	Pouch	LCO (LNCO)	x	x	2p, 3p	9	Cable 0.3 [mΩ]	0.08	x			x	x					x	x	x					
Fill	[16]	2021	Pouch	LCO (LNCO)	x	x	2p	2	Cable 0.3 [mΩ]	0.08						x				x	x	x					
Fill	[57]	2022	Pouch	LCO (LNCO)	x	x	2p, 3p, 4p	15	Cable 0.3 [mΩ]	0.08	x			x	x	x				x	x	x	x				
He	[52]	2023	N.A.	N.A.		x	2p	6	N.A.	N.A.	x				x					x		x					
Hosseinzadeh	[50]	2021			x				N.A.	N.A.	x[a]	x[a]	x[a]		x[a]			x[a]	x[a]			x[a]	x[a]	x[a]			
Jocher	[18]	2021	18 650	NMC	x	x	2p	4	Shunt 1 [mΩ]	1.2		x										x					
Li C.	[20]	2022	18 650	NMC	x	x	2p, 4p	4	Hall	0.2, 7		x	x					x[a]	x[a]	x		x	x	x		x[a]	
Li Z.	[54]	2022	26 650	LFP	x	x	2p	2	Shunt 1 [mΩ]	1.3	x[a]		x		x[a]	x[a]				x[a]		x		x			
Liu	[34]	2019	Pouch (High-Power)	NMC-LCO		x	6p	6	Shunt 10 [mΩ]	1	x		x[a]							x		x				x[a]	x[a]
Luan	[23]	2021	18 650	NMC	x	x	2p	2	Hall	N.A.	x[a]				x[a]				x[a]			x			x[a]		
Luca	[58]	2021	18 650	NMC		x	8p, 9p	9	Hall	1.42								x	x	x		x	x		x		
Marlow	[30]	2023	Pouch (High-Power)	NMC-LCO	x	x	2p	12	Shunt 1 [mΩ]	1			x	x								x	x			x	x
Reiter A.	[56]	2023	Pouch	NCM		x	14s2p	28	Not meas.	N.A.																	
Reiter C.	[9]	2019	18 650	NCA	x	x	2p	2	Hall	N.A.	x											x					
Rumpf	[49]	2018	26 650	LFP	x				Not meas.	N.A.	x[a]	x[a]	x[a]		x[a]				x[a]			x[a]		x[a]			
Schindler	[39]	2021	18 650	NMC	x	x	2p	28	Shunt 1 [mΩ]	1.2	x				x					x		x				x	x
Tian	[26]	2022	N.A.	LFP, NMC	x	x	2p, 3p, 4p	27	Hall	N.A.	x				x		x					x					
Wang	[33]	2022	Pouch	NMC	x	x	2p	2	Hall	N.A.	x									x		x		x[a]			
Wang	[40]	2019	18 650	LCO	x	x	2p, 2s2p, 2p2s	12	N.A.	N.A.	x				x				x			x	x			x	x
Ye	[61]	2019	18 650	NCA	x	x	2p4s, 4s2p	8	Not meas.	N.A.					x				x						x		
Zhang Y.	[35]	2018	Pouch	LFP	N.A.	x	5p	5	Hall	N.A.										x		x					

[a] Experimental campaign not performed.

necessity of cable sizing trade-offs. Signal detectability increases with higher resistance, while the reduction of cable resistance mitigates the impact of measurement techniques on the signal. Hall-effect sensors have undergone significant improvements over time and are commonly favoured in research that places a high value on the impact of the test bench on the results, given their inclusion in an external circuit [58].

In scenarios where the system exhibits complexity and involves multiple variables, it is imperative to establish a structured and controlled experimental design to guarantee the precision and dependability of outcomes. The recent interest in implementation of Design of Experiments (DoE) in the battery community is attributed to its capability to identify the most significant factors in a system, resulting in cost savings through reduced experimentation time and resources [63]. Previous applications encompass various areas such as the ageing of individual cells, energy capacity, electrode formulation and material synthesis, thermal design, and charging [64]. Some studies are also available on pack-level thermal design, with optimisation of the cooling mass properties [65] and cooling design structure [66]. In addition, the empirical data acquired through DoE can be employed to define parameters for theoretical models. Nevertheless, the empirical models are usually confined to linear or quadratic relationships, which may not be sufficient to describe the effects of individual cell properties and module features on system imbalances. The results obtained from DoEs are commonly subjected to statistical techniques such as Analysis of Variance (ANOVA) and graphical investigation for analysis [67]. Recently, researchers have focussed on leveraging the accuracy of Machine Learning (ML) modelling techniques for the identification of the feature-response relationships [68,69]. Despite exhibiting robust predictive capabilities, ML models are commonly acknowledged to have limitations in terms of "transparency" [70]. This can pose a challenge in "screening" type studies [71]. To this end, Explainable Machine Learning (XML) algorithms have been established as effective Artificial Intelligence (AI) techniques [72]. These algorithms provide various metrics for questioning the ML models decisions and enhancing their interpretability. Faraji et al. [73,74] proposed a systematic methodology for the analysis of the impact of manufacturing process on the electrochemical properties and performance of cells. These studies utilise XML techniques such as Mean Decrease in Impurity (MDI), Shapley values and Accumulated Local Effects (ALE) to derive insights on the importance of high-volume manufacturing features and their contributions to cells' Key Performance Indicator (KPI).

The examination of the literature exposes certain gaps. The preponderance of testing at the module level was executed on a pair of cells that were connected in parallel. While crucial for comprehending the sources of imbalances in the module, these factors often encompass magnitudes that exceed those encountered in practical scenarios involving a greater quantity of cells. The limited size of the parallel-connection studies batches hinders the statistical validity of variations in cell characteristics, thereby posing a challenge to the generalisation of findings due to the absence of a link with the typical distributions of freshly produced cells. Few precise studies emphasise the importance of controlling the impact of the experimental environment on phenomena occurring at the module level. Despite being widely recognised as a critical factor in the non-uniform performance of the cells within modules, the mapping and assessment of electrical connection resistances are infrequently conducted, which is a topic of controversy [75]. The employment of Hall effect sensors for current monitoring in experimental research was deferred until recent years owing to their inadequate precision, and instead, shunt-type sensors were favoured. The capacity of Hall-type sensors to integrate with an external circuit without disrupting the current distribution in the string, coupled with their progressively improving precision, are crucial factors in detecting the origins of heterogeneity that will not depend on the experimental setup. Notwithstanding the growing interest in DoE, previous research has solely concentrated on its application at the level of individual cells. The utilisation of the DoE methodology in investigating modules and packs is a crucial aspect as it pertains more closely to practical battery applications. Despite its potential to enhance analytical capabilities, the utilisation of emerging ML modelling and transparency-enhancing approaches to investigate modules and packs feature importance has been limited. To enable such investigations, this section aimed to address the dearth of comprehensive data that is often

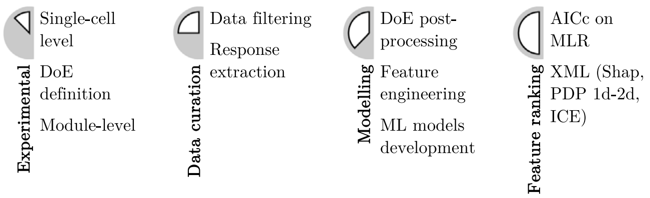

Fig. 1. Analysis flowchart: principal stages and methods involved.

encountered in the literature, which may be attributed to the limited scope of experimental studies that focus on cell pairs or the absence of information pertaining to the involved cells and experimental setups.

The ability to anticipate anomalous behaviours of parallel cells is crucial to ensuring the longevity and safety of battery packs. It is necessary to understand the predominant factors contributing to the inconsistent performance of the module and leverage this information to create predictive models for system optimisation. To the best of the authors' knowledge, there have been no previous attempts to apply a DoE at a module level for purposes other than thermal design optimisation [65,66]. Furthermore, there has been a lack of thorough utilisation of XML techniques to identify and prioritise the influence of features on battery systems imbalanced performance. Following the aforementioned literature review and gap analysis, the novelty and contributions of this study can be summarised as follows:

- **Experimentally investigate the impact of individual cell and module properties on the performance of parallel connected cells:** A parallel module-level experimental campaign is for the first time conducted using the concepts of a full-factorial DoE, allowing to consider all features independently and thereby discern their individual impact and interactions among control variables.
- **Identify the most influential elements in modules' imbalanced performance:** The DoE campaign ensures results' statistical relevance and allows to methodically isolate and rank the impact of key factors such as interconnection resistance, operating temperature, cells chemistry combination and ageing impact on parallel strings current, temperature and time-to-self balance deviations.
- **Combine novel interpretable machine learning techniques with established linear regression strategy:** Novel XML techniques together with established statistical analysis are applied to increase models predictability by capturing non-linear relationships and elevate their interpretability. Neural Networks and Random Forest models are trained and optimised to get alternative information to multivariate linear regression, outlining their benefits and drawbacks in unveiling parallel connected cells heterogeneous performance contributors.

The article is organised as follows. Section 2 presents the methodology and developed framework for testing analysis and feature importance derivation. The experimental strategy and practical details are outlined in Section 3 which extends from the individual cell characterisation campaign through to the test bench verification and the execution of the DoE. Section 4 derives a multivariate linear regression model for experimental data analysis and identifies the major factors influencing the modules' imbalanced performance. Following that, in Section 4, significant ML models are compared and submitted to interpretability-expanding techniques to validate their benefits and limitations over traditional statistical approaches. Finally, in Section 5, the most important findings, conclusions and future research directions are pointed out.

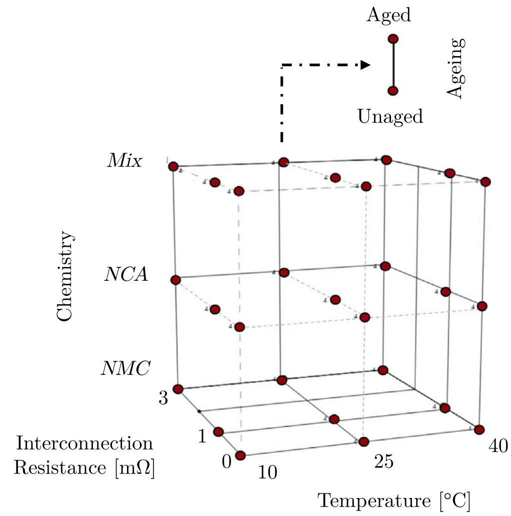

Fig. 2. Graphical representation resulting from the four factors and respective levels of the full-factorial DoE.

2. Methodology

The process of determining and prioritising the factors that contribute to the uneven performance of modules involves a series of steps, which are reported in Fig. 1. The initial step entails the preparation and execution of a DoE campaign. Step 2 is about data curation, filtering and management while the subsequent procedures facilitate the attainment of two distinct methodologies for conducting feature analysis. In Step 3, the DoE results are post-processed and the statistical and ML models are trained. Step 4 focuses on enhancing the models' interpretability. The first feature analysis approach relies on conventional multivariate linear regression, which involves minimising the Akaike Information Criterion (AICc) to reduce the number of variables [76]. The second approach aims to enhance the interpretability of machine learning models by leveraging XML techniques while capitalising on their accuracy.

2.1. Design of experiments

Various areas of study in DoEs have been documented in the literature [64], with this paper falling under the "screening" definition. A screening or characterisation study aims to identify and rank the relevant features to the interested variables and responses. Technically, a rich and representative dataset is first needed for a comprehensive study of the influence of the multiple elements on the uneven performance of modules. There are several distinct stages involved in a DoE [77]. First, the control parameters, as well as their ranges and levels, are chosen based on the literature, resulting in a cube such as the one reported in Fig. 2 for visualisation purposes. According to the expert's view and literature review, the interconnection resistance could extend up to a 2.5% ratio to the energy cell's internal resistance [78]. In this study, it ranges from 0, 1 to 3 mΩ to include poor busbar design scenarios, as detailed in Table 2. The testing temperatures (10 °C, 25 °C, 40 °C) are selected to match the typical laboratory temperatures and hence most comparable to prior and future investigations as well as real-world operating conditions [79]. The choice of NMC and NCA chemistries is based on their comparable characteristics and potential suitability for re-purposing in second-life applications [25]. A "Mix" chemistry configuration refers to a combination of two NMC and two NCA cells in the parallel string being tested. Although mixing chemistries is an unconventional approach, recent publications are exploring hybrid configurations of power and energy cells [80]. The

Table 2
List of the control and response variables included in the study. N_p refers to the number of cells connected in parallel (in this study, $N_p = 4$). I_{Cellk} denotes the current delivered by the kth cell, I_{Mod} is the module input current. T_{Amb} is the ambient temperature, SoC_{Cellk} and T^{Surf}_{Cellk} are the SoC and surface temperature of the kth cell, respectively. t_1, t_2, and t_{end} are the time instants used to split the cell current distribution, as depicted in Fig. 3.

Control variables	Levels	Response variables		Unit
Interconnection resistance [mΩ]	[0, 1, 3]	$\sigma I_{Start} = \frac{1}{t_1}\int_0^{t_1}\left(\sqrt{\frac{1}{N_p-1}\sum_{k=1}^{N_p}(I_{Cellk} - I_{Mod}/N_p)^2}\right)dt$	(1)	[A]
Temperature [°C]	[10, 25, 40]	$\sigma I_{Mid} = \frac{1}{t_2-t_1}\int_{t_1}^{t_2}\left(\sqrt{\frac{1}{N_p-1}\sum_{k=1}^{N_p}(I_{Cellk} - I_{Mod}/N_p)^2}\right)dt$	(2)	[A]
Chemistry [–]	[NMC, NCA, Mix]	$\sigma I_{End} = \frac{1}{t_{End}-t_2}\int_{t_2}^{t_{End}}\left(\sqrt{\frac{1}{N_p-1}\sum_{k=1}^{N_p}(I_{Cellk} - I_{Mod}/N_p)^2}\right)dt$	(3)	[A]
Ageing [–]	[Aged, Unaged]	$\Delta SoC_{Max} = \max(SoC_{Cellk}) - \min(SoC_{Cellk})$	(4)	[%]
		$\Delta SoC_{End} = \max(SoC_{Cellk}\|_{t=t_{End}}) - \min(SoC_{Cellk}\|_{t=t_{End}})$	(5)	[%]
		$\Delta T^{Max}_{Net} = \max(T^{Surf}_{Cellk} - T_{Amb}) - \min(T^{Surf}_{Cellk} - T_{Amb})$	(6)	[°C]
		$\sigma T_{Mean} = \frac{1}{t_{End}-t_1}\int_{t_1}^{t_{End}}\left(\sqrt{\frac{1}{N_p-1}\sum_{k=1}^{N_p}(T^{Surf}_{Cellk} - T_{Amb})^2}\right)dt$	(7)	[°C]
		TTSB		[s]

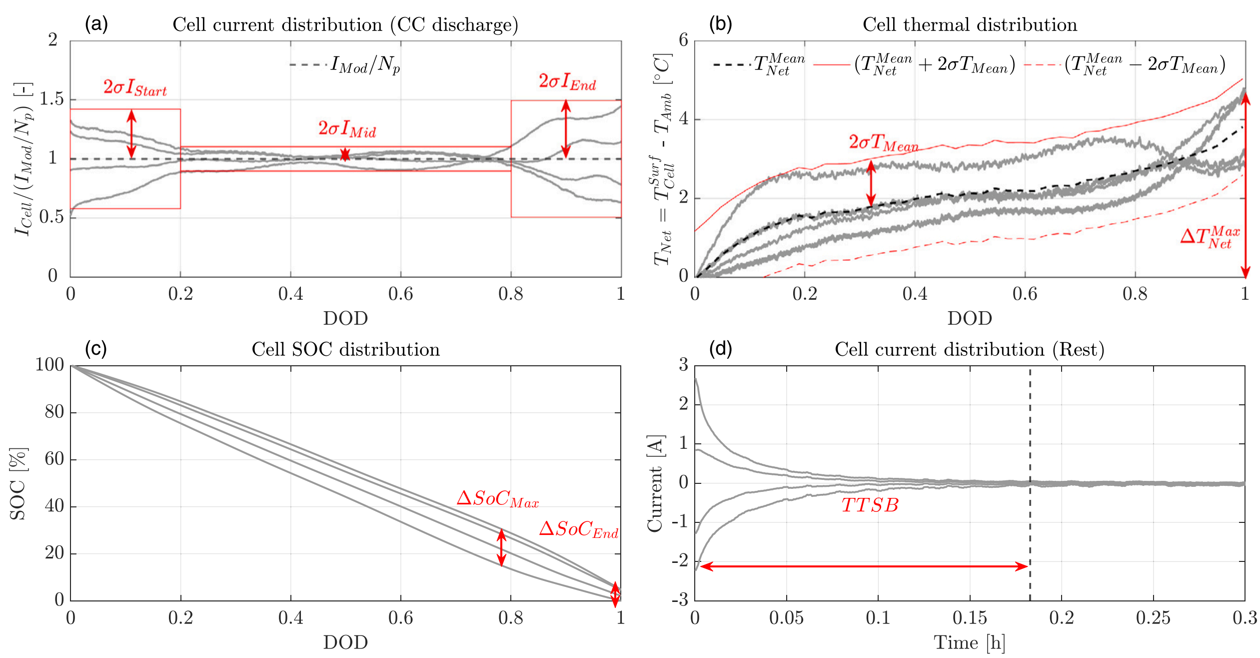

Fig. 3. Visualisation of the eight response variables (in red) extracted in this study from the experimental results (in grey) for one of the 54 performed experiments. (a) Current, (b) temperature, (c) SoC related responses, (d) TTSB. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

rationale behind this approach is to combine the advantages of both chemistries, i.e., high power and high energy density, respectively. In this context, the proposed methodology is a first step towards the development of a framework that can be used to assess the performance of hybrid configurations. Finally, the addition of a cell with low State of Health (SoH) is done in order to gain a better knowledge of the behaviour of cells when there is either a failed contact or a high resistance and capacity gradient present.

A set of eight response variables grouped in four distinct categories is chosen based on the specific phenomena under analysis in this study and are reported in Table 2. Current distribution entails the first category. As shown in Fig. 3, there are three stages to the current propagation among the four cells during discharge. The first is convergent, the second is stable, and the third is divergent. Depending on the stage, the current average standard deviation at the beginning, middle and end of cycle ($\sigma I_{Start}, \sigma I_{Mid}, \sigma I_{End}$) are calculated as in (1), (2) and (3), respectively. The second group is closely related to the first one and gives information about the maximum absolute and the end difference in State of Charge (SoC) of cells (ΔSoC_{Max}, ΔSoC_{End} computed as shown in (4) and (5), respectively). The SoC of each individual cell is determined utilising Coulomb counting [81]. The third variable category pertains to surface temperatures. The symbols ΔT_{Max} and σT_{Mean}, calculated according to (6) and (7), denote the maximum temperature increase and the mean standard deviation of the temperature gradient across cells throughout the discharge process, respectively. The Time To Self Balance (TTSB) characterises the time taken (measured in seconds) by the cumulative balancing currents to settle down to 200 mA during the post-discharge period. Fig. 3 provides a visual representation of the significance of the eight response variables.

Subsequently, the module experimental design is chosen. The selection of a "Full-Factorial" DoE is based on the fact that it is a comprehensive methodology that entails the examination of all conceivable combinations of the four factors and their respective levels, leading to a total of 54 tests. To mitigate the impact of developing changes in individual cells characteristics over multiple cycles, a randomised sampling methodology is employed. The methodology entails the random selection of groups comprising four cells from the newly available batches. This means that at every test run four cell numbers (1–20) and module locations (1–4) are randomly allocated. In this way, it is not the same set of cells being tested throughout the entire campaign. By decreasing the number of cycles that each cell undergoes, it is feasible to mitigate the impact of untraceable ageing on the tests. The initial 30 experiments are incorporated within an I-optimal design [82], which is formulated using precise objective optimality criteria through the utilisation of Design Expert software [83]. The primary benefit of this methodology lies in the potential to construct an initial model. The inclusion of an intermediary stage can prove advantageous in validating the efficacy of the process and managing any potential errors. In addition, optimal designs have the potential to yield effective models with reduced experimentation, provided that their outcomes align favourably with those of the full-factorial DoE. Owing to the significance and magnitude of this particular area of investigation, it has been excluded from the objectives of the present study and reserved for subsequent research.

The DoE procedure encompassing the execution of experiments is delineated in Section 3. The data collected is analysed using both traditional statistical models such as multivariate linear regression based on DoE and modern machine learning models like Neural Networks and Random Forests. The ultimate objective of the screening investigation is to ascertain and prioritise the effect of included features. The control variables are subject to screening through either AICc or ranking and visualisation using XML techniques, depending on the selected empirical model.

2.2. Modelling

The aim of conducting a systematic DoE is to acquire a comprehensive empirical model that establishes and unveils a relationship between the control and response variables, denoted as x and y, respectively. Prior to the application of any modelling methodology, it is preferred to first perform feature scaling. This particular stage holds significant importance for both conventional empirical and machine learning models. When employing polynomial models, the process of scaling the input features enables a direct comparison of the coefficients (β). Through the modelling process, the coefficients can be ranked based on their magnitude. To reduce input sensitivity in ML models, feature scaling is fundamental. The normalisation of features in this study involves the removal of the mean and scaling to unit variance. This is achieved through the use of the formula:

$$z = \frac{(x-\mu)}{\sigma} \tag{8}$$

where μ represents the feature mean and σ represents its standard deviation.

Multivariate linear models are conventionally derived from full factorial DoE. The MLR model is selected for its transparency and ability to facilitate interpretability of input–output relationships in a coherent manner. This is achieved by enabling a fair comparison of the considered features through model weights. The present investigation incorporates main effects, two-way interaction terms, and second-order polynomials to account for potential curvature when modelling the responses. The exclusion of higher orders polynomials has different reasons, including the sensitiveness to the order of the polynomial, the tendency to overfit and the worsening of interpretability. The model is presented in the form:

$$\hat{y} = \beta_0 + \sum_{i=1}^{k} \beta_i z_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} z_i z_j + \sum_{i=1}^{k} \beta_{ii} z_i^2 + \epsilon \tag{9}$$

where k is the number of control variables, z_i refers to the ith normalised control variable (with $i = 1, \dots, k$), and ϵ represents the random error. In total, there are 14 coefficients, in addition to the intercept or bias (β_0). The estimation of β parameters is achieved through the least-squares method. MLR models perform well in linear spaces and can handle a large number of features. Nevertheless, they show decreasing performance when the signal to noise ratio is high or the underlying function is not truly linear or non-monotonic. To unveil non-linear relationships, ML models such as RF and NN allow to approximate more complex shapes, to the detriment of features importance understanding. As it is shown in Section 4, NN responses can be more interpretable than RF models and closer to the MLR ones. Besides, NN models present a larger set of tuneable hyperparameters, which can improve the predictability of the models and capture strongly non-linear relationships. Consequently, NN are included in this paper as they can play a junction role between the interpretability of MLR and the predictability of RF models. Despite that, they are still prone to overfitting and commonly addressed as black-box models. Hence, the nature of variable interactions and the high level features learned by the network require XML methodologies to be discernible. Further details on the selected ML models are offered in Appendix A.

To establish whether linear regression is sufficient in the evaluated case or more complex models are required to unveil the input–output relationships, the three different approaches are benchmarked. To evaluate the performance of the models and the accuracy of the representations, here the metrics of coefficient of determination (R^2) and Root Mean Squared Error (RMSE) are computed according to Eqs. (10) and (11), respectively.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \frac{1}{N}\sum_{j=1}^{N} y_j)^2} \tag{10}$$

$$RMSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{11}$$

R^2 and $RMSE$ are statistical metrics commonly used in regression analysis, where N is the number of observations. R^2 quantifies the proportion of variance in the dependent variable that can be accounted for by the independent variable(s) in the model. $RMSE$ measures the standard deviation of the residuals. Given that the dataset under investigation for this study exhibits a restricted number of samples, a reliable validation strategy is critical. For this purpose, the K-fold Cross Validation (CV) approach is utilised to mitigate estimation bias and yield a more generalisable assessment of model accuracy. While the K in K-fold algorithm could be any values between 2 and the number of samples minus 1, here without the loss of generality, the value of K is chosen to be 5 to ensure that the populations are of sufficient size for both the training and testing (validation) phases. This means, the dataset is partitioned into a ratio of 80% for training and 20% for validation-testing, and the procedure is applied to all five groups upon random selection of samples.

2.3. Feature importance

The present study employs a statistical technique to restrict the inclusion of linear, interaction and quadratic terms in the MLR model, whereby the corrected AICc is minimised [76].

$$AIC(M_k) = -2\log L(M_k) + 2k \tag{12}$$

$$AICc(M_k) = AIC(M_k) + \frac{(2k^2 + 2k)}{(n-k-1)} \tag{13}$$

The logarithm of the likelihood function, denoted as $\log L(M_k)$, pertains to the model M_k. The variable k represents the number of features incorporated in the model, while n denotes the size of the sample. This methodology involves the reduction of the control variables to only statistically significant ones. Then, the process of feature scaling serves to ensure comparability of coefficients and enhance interpretability. In fact, in this case the MLR model feature importance analysis can be performed via its resulting weights.

Opaque machine learning models necessitate alternative methodologies for determining feature ranking, as traditional methods are not applicable [71]. Prior studies have delineated the present Explainable Artificial Intelligence (XAI) methodologies that are appropriate for providing post-hoc interpretability for opaque models [84]. The Model-agnostic XML techniques are not reliant on any specific model and are intended to have broad applicability, in contrast to Model-specific techniques. Their operations are solely based on the relationship between the control and response variables. The present investigation employs SHapley Additive exPlanations (SHAP) [85] as the method of choice for identifying feature relevance. Additionally, the study utilises the capabilities of Partial Dependence Plot (PDP) [86] and Individual Conditional Expectation (ICE) [87] plots for visualising feature impact.

Lundberg and Lee [85] proposed the SHAP technique, which utilises game theory-derived Shapley values to provide an explanation for individual predictions. The Shapley value can be defined as the mean level of contribution that a given member provides to the overall value of the coalition, taking into account all feasible permutations [88]. Given the requirement of theoretically sampling the coalition values for every possible feature permutation, the model must undergo an equivalent number of evaluations. As the number of features increases, there is a corresponding rise in computational effort. Lundberg and Lee [85] propose the utilisation of the Shapley Kernel as a means to address the aforementioned constraint, thereby enabling the estimation of Shapley values with a significantly reduced number of samples. However, in our scenario with four input features it is possible to quickly sample the resulting 64 coalitions. The explanation can be defined as:

$$g(z') = \phi_0 + \sum_{j=1}^{M} \phi_j z'_j \tag{14}$$

where g is the explanatory model, z' is the simplified features, M is the maximum coalition size and ϕ_j are the Shapley values for a feature j, included in the following equation.

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f_{S\cup\{j\}}(x_{S\cup\{j\}}) - f_S(x_S)] \tag{15}$$

Following from game theory, in Eq. (15) F is the number of features used in the model, S is a subset of the features, $x_{S,S\cup\{j\}}$ are the vector of feature values of the instance to be explained and $f(\cdot)$ is the value function. The present study is concerned with the use of Shapley values as a means of quantifying variable importance, which has been previously demonstrated in research as an alternative to the conventional functional ANOVA approach [89,90]. However, SHAP does not quantify the significance of a particular feature in the actual world; rather, it measures the importance of a feature to the model. Therefore, the process of extrapolating beyond the confines of the given space is intricate. The SHAP method facilitates the identification of feature importance, however, it exhibits limitations in terms of visualising the distribution of feature impacts across the evaluated space. This study incorporates ICE plots and PDPs to facilitate the visualisation of feature importance. ICE plots depict a single line per instance, illustrating the alteration in the instance's forecast as a feature is modified while the remaining features remain constant at their observed values [87]. The use of PDPs enables the assessment of the impact of one or two specific features on the anticipated outcome of a ML model, while controlling for the average effect of the remaining features [86]. Mathematically, they can be expressed as:

$$PD_{X_S}(x_S) \overset{\text{def}}{=} \mathbb{E}_{X_C}[f(x_S, X_C)] = \int f(x_S, x_C) p(x_C) dx_C \tag{16}$$

The function $f(x_S, x_C)$ represents the response of a given set of samples, x_S and x_C, for the features in X_S and X_C, respectively. The PDPs and ICE plots are connected through the evaluation of the response function $f(x_S, x_C(i))$ at x_S, which defines each individual line. According to [72], the use of PDPs enables the visualisation of the feature of interest and response variable are related, thereby facilitating the identification of whether the relationship is linear, monotonic, or more intricate in nature. However, the principle of PDPs based on averages does not necessarily eliminate the possibility of interference from interacting variables. This limitation is circumvented by the complementary character of PDPs and ICE plots, which perform optimally when combined as in the presented research.

Table 3
Technical specifications of the tested cells [94,95].

Manufacturer	LG Chem	Samsung
Model	INR21700-M50T	INR21700-50E
Positive electrode	$Li(NiCoMn)O_2$	$Li(NiCoAl)O_2$
Negative electrode	Graphite and silicon	
Size (diameter × length)	21.44 × 70.80 mm	21.25 × 70.80 mm
Weight	69.25 g	69.00 g
Nominal capacity (C_n)	4.85 Ah	4.90 Ah
Nominal voltage	3.63 V	
Charge cutoff voltage	4.2 V	
Discharge cutoff voltage	2.5 V	
Cutoff current	50 mA	

3. Experimental study

The experimental configuration employed in this study comprises a battery cycler (Arbin Instruments LBT22013), Hall effect sensors (Honeywell SS495A), an external circuit for the 5 V power supply and a self-constructed cell holder. The experimental setup thermal chamber (Amerex IC500R) is used to test at varying temperatures, specifically 10 °C, 25 °C, and 40 °C. All measurements are transmitted to a designated host computer equipped with the MITS Pro software. The data obtained from an auxiliary measurement system that records signals from thermocouples and voltage sensors are transmitted through a Controller Area Network (CAN) communication protocol. Further details on the Stanford Energy Control Laboratory testing equipment can be found in [91,92]. The research endeavour is divided into two distinct segments. The initial focus of the experiment pertains to the assessment of the distributions of individual cell parameters to isolate their contribution to the parallel connection performance and to identify eventual outliers. The second part is focused on capturing the performance of the module by developing the experimental setup and limiting its influence on the results of the DoE campaign.

3.1. Single cells' characterisation

A total of 40 cells are characterised independently by connecting them individually to the channels of the battery cycler at its lowest and most precise range. Two sets of 20 newly manufactured LGM50T and Samsung50E cells are characterised. The technical specifications of the tested cells are reported in Table 3. Both LG and Samsung cells utilise Silicon-doped graphite (SiC) based negative electrodes. The positive electrode of LG cells consists of Nichel Manganese Cobalt (NMC) 811 oxide, whereas Samsung cells utilise Nichel Cobalt Aluminium (NCA) oxide as their cathode material. The derivation of cell-to-cell variability is carried out to enhance the interpretability of any module-level imbalance that may arise due to cells' heterogeneity. The protocol for measuring capacity involves a charge of Constant Current Constant Voltage (CC-CV) at a rate of C/3 until a cut-off current of 50 mA is attained at 4.2 V. This is followed by a Pseudo-OCV discharge at a rate of C/20 until cut-off at 2.5 V. The procedure for characterising ohmic resistance involves a subsequent charge with identical properties to the initial charge, succeeded by discharges at 10% state of charge at a rate of C/3, and a resting period of one hour to reach a state of equilibrium. What follows is a Hybrid Pulse Power Characterisation (HPPC) profile, characterised by a charge/discharge ratio of 0.75 and pulse durations of 10 s, as per automotive standard [93]. The experiments at the single-cell level are conducted at a temperature of 23 °C, with results reported in Table 4.

Table 4
Single cell campaign capacity and internal resistance measurement results.

Variable	Unit	NMC			NCA		
		μ_x	σ_x	$\frac{\sigma_x}{\mu_x}$ [%]	μ_x	σ_x	$\frac{\sigma_x}{\mu_x}$ [%]
Temperature	[°C]	23.20	0.47	2.03	23.22	0.41	1.77
Capacity	[Ah]	4.86	0.03	0.68	4.96	0.01	0.28
Ohmic Res.	[mΩ]	26.95	0.65	2.41	19.32	0.49	2.53
90% SoC	[mΩ]	26.74	0.59	2.20	18.96	0.38	2.00
80% SoC	[mΩ]	26.38	0.58	2.19	18.77	0.45	2.42
70% SoC	[mΩ]	26.30	0.58	2.19	18.71	0.42	2.23
60% SoC	[mΩ]	26.36	0.58	2.22	18.84	0.44	2.32
50% SoC	[mΩ]	26.43	0.61	2.29	18.89	0.45	2.39
40% SoC	[mΩ]	26.56	0.63	2.35	18.86	0.47	2.47
30% SoC	[mΩ]	26.82	0.65	2.43	19.14	0.50	2.61
20% SoC	[mΩ]	27.13	0.68	2.49	19.71	0.55	2.79
10% SoC	[mΩ]	29.84	0.98	3.27	21.97	0.73	3.34

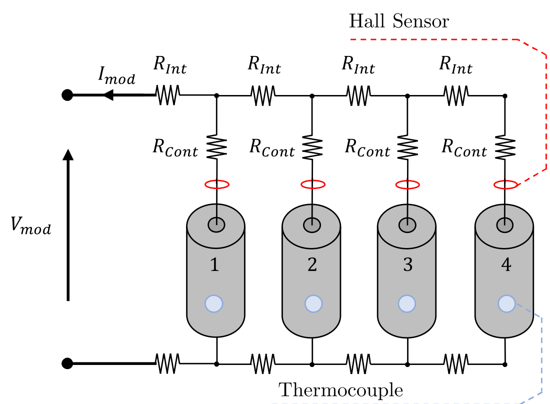

Fig. 4. Circuit diagram of the module test bench with implemented sensors locations.

Table 5
Experimental setup contact and interconnection resistances measurement results.

Item	μ [mΩ]	σ [mΩ]
$R_{Int\rightarrow(Full-busbar)}$	$5.4e^{-3}$	$8.18e^{-4}$
$R_{Int\rightarrow(Shunt=1\ m\Omega)}$	1.05	0.08
$R_{Int\rightarrow(Shunt=3\ m\Omega)}$	3.02	0.05
$R_{Contact}$	1.21	0.04

3.2. Module-level measurement setup

The experimental campaign required to minimise the impact of the setup on the current distribution. Initially, the extent of the interconnection resistances of the three busbar types manufactured is established by measuring via shunt-type resistors the voltage drop across each segment. Subsequently, the contact resistance amid the cell poles and the experimental arrangement is evaluated at progressively elevated levels of pressure. This is achieved by tracking the torque applied to the bolt-type connections via a dynamometric wrench and measuring the resulting change in voltage drop and hence resistance values. The critical aspect that is mapped pertained to the variability between each of the four contact points, as they are in series with the cells. The applied torque is chosen to achieve both minimum contact resistance within the branch and minimum oscillation among the branches. Table 5 presents a summary of the outcomes obtained from the mapping of interconnection and contact resistances.

The literature indicates that the performance of the module is influenced by the positioning of the cells within the string. Consequently, the decisions pertaining to the placement of the aged cell and the combination of NMC and NCA are approached from a worst-case scenario standpoint. This means ensuring the largest possible contribution to the string performance imbalance. In the DOE at the “Aged” levels, one old cell is positioned at the farthest distance from the terminals. The “Mix” level comprises a pair of NCA cells situated in proximity to the terminals, with the remaining two positions occupied by NMC cells. Likewise, the decision to utilise the ladder connection is also undertaken to pinpoint the most unfavourable operational circumstances for a 4P configuration. The obtained results can be directly transferred to other architectures, such as the “Z” one, which are commonly showing lower imbalance levels. The measurement of current is conducted through the utilisation of Hall effect sensors that are strategically positioned within the negative pole of every cell. Fig. 4 presents the electrical circuit diagram of the experimental setup, including the tracked resistances and the Hall sensors locations. The Hall sensors are embedded within ferrite rings to facilitate the transmission of their magnetic field and enhance their signal. The stability of the signal is achieved by means of connectors’ upward and downward filters. Before starting the campaign, every Hall sensor underwent calibration procedures utilising established current values in order to reconstruct the mapping of voltage and current curves. The low sensitivity drift of 0.05[%/K] guarantees that there is minimal interference of temperature in the measurements. The internal circuit of the cycler is utilised to measure the terminal voltage. Thermocouples are utilised to measure the temperature of each cell’s middle surface, in conjunction with an extra sensor to log ambient conditions. The module-level testing procedure entails a CCCV charging method at a rate of C/3 until the cut-off current of 50 mA is reached at 4.2 V, followed by a discharge at a rate of 0.75C.

4. Results and discussion

This section describes the outcomes of the experimental campaigns conducted and presents the statistical analysis of the variables under assessment. Initially, the data from the experimental campaign at the single-cell level is introduced to describe the parameters’ distribution of the 40 cells. The results of the experimental campaign at the module level are utilised to train and test three models. Subsequently, the models are statistically analysed using specific methodologies to obtain the most influential features affecting the imbalanced performance of parallel-connected cells. All experimental findings are found in the associated article in Data in Brief [96].

The experimental campaign on individual cells aims to define the distribution of internal characteristics of cells once they have exited the manufacturing process. The 40 observations of nominal capacity and internal resistance are obtained at an operating temperature of 23 °C and are presented in Fig. 5. As depicted in Fig. 5(a), the pseudo OCV discharge demonstrates a nominal capacity consistent with the nominal value declared by the cell manufacturers. NMC cells exhibit a distribution that is not in line with a normal bell curve. A detailed analysis of the 20 NMC cells led to the identification of 3 outliers originating from different batches and thus excluded from the subsequent campaign. The reasoning regarding NCA cells is diverse, as they are normally split around the mean. As stated in Table 4, the ratio between the standard deviation and the mean of NMC cells is approximately twice that of NCA cells. The distributions of cell parameters in terms of internal resistance are more uniform. The average internal resistance of NCA cells is approximately two-thirds that of NMC cells. The increase in standard deviation is significant as the SoC decreases. The impact of this phenomenon on the performance of parallel strings at low SoC may serve as a basis for future investigations.

By characterising the available cells population it is possible to carry more information on the causes of imbalances when scaling up to module level for the second experimental campaign. DoE approach results enable the derivation of multivariate linear equations describing the tested space. These equations can then be displayed via contour plots, as in the case of Fig. 6 including three out of the eight response variables of the MLR model, namely σI_{Start}, σT_{Mean}, and $TTSB$. The x

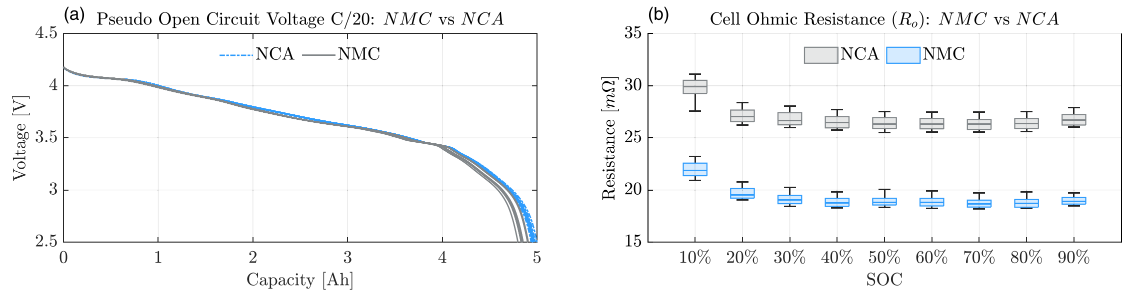

Fig. 5. (a) Pseudo OCV curves of the fresh 40 cells under C/20 discharge test procedure. (b) Boxplot of cells internal resistances at 10% SoC intervals.

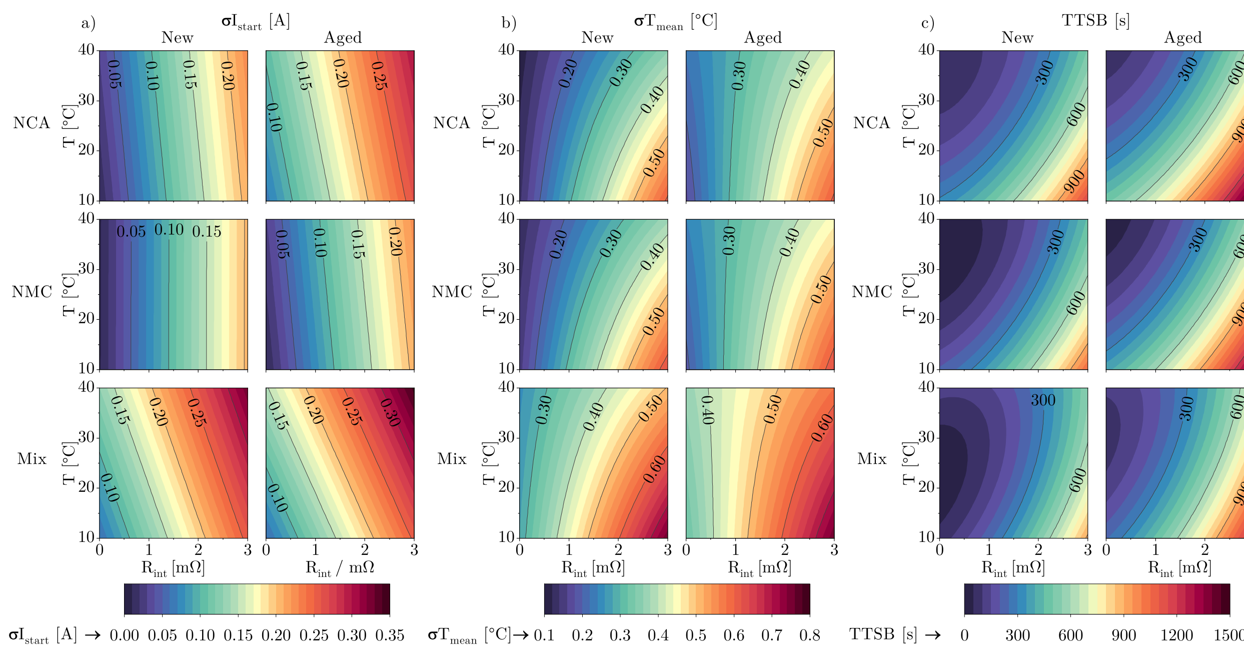

Fig. 6. Contour plots derived from the multivariate linear models simulated space for (a) σI_{Start}, (b) σT_{Mean}, (c) $TTSB$. The horizontal and vertical axes of individual contour plots are the interconnection resistance and the test temperature, respectively. The columns of each group of six contour plots have new and aged cells configurations, respectively, while the rows differ upon chemistry (NMC, NCA, Mix). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and y axes of the subgraphs stand for the interconnection resistance and test temperature, respectively. The columns indicate instances where all four cells are new or one of them is aged. The rows comprise NCA, NMC, and NCA-NMC mixed chemistry tests, from top to bottom. Contour plots are found to be effective in providing information on the variation of response variables depending on control variables and in classifying features according to their importance. A common trend observed in all three graphs is that as the interconnection resistance increases, the performance of the cells connected in parallel deteriorates. These performances can be classified into four main categories. The main categories considered in the response variables are current imbalances, SoC, temperature, and lengthening of the TTSB. In the case of Fig. 6(a), an imbalanced performance implies that the four cells do not exhibit an equal level of current flow. As the interconnection resistance increases, the σI_{Start} rises as well, up to over 300 mA in the worst case scenario. It is evident that while temperature has a minimal impact on individual NCA and NMC cells, it plays a dominant role in the case of combinations of cells with different chemistries. In the bottom row of Fig. 6(a), diagonal lines are observed, indicating an almost equal relevance between interconnection resistance and operating temperature. As the temperature increases, it is observed that the initial current distribution deteriorates, suggesting a decreasing equilibrium between the current of each cell. If the operating temperature is raised from 10 °C to 40 °C, there is a resulting difference of 70 mA at a constant interconnection resistance. If an aged cell is inserted, the current on the other cells is negatively affected. The discussion regarding the σT_{Mean} in Fig. 6(b) differs. The operating temperature assumes a more significant role. The relationship between the operating temperature and interconnection resistance is non-linear and exhibits curvature. As the operating temperature increases, the cells operate under more uniform conditions. The reduction of the gradient can reach 0.1 °C in cases of interconnection resistance of 1 mΩ, up to 0.2 °C in the case of 3 mΩ at the tested operating temperature extremes. The NMC and NCA cells exhibit similar and comparable characteristics under the tested conditions. The oscillation range of the gradient is in the order of 0.1–0.6 °C for both, as the inputs vary. If the two chemistries are mixed, the temperature difference between the cells increases to over 0.7 °C. Although the introduction of an aged cell does not significantly impact the temperature distribution in the case of single chemistry, this changes when multiple chemistries are mixed. In such cases, the average temperature gradient increases significantly, and the gradient range reaches 0.35–0.75 °C. Fig. 6(c) illustrates a significant curvature of the TTSB between interconnection resistance and operating temperature. It can be observed that the latter two control variables have the greatest influence on TTSB, with ageing and cell mixture playing a less significant role. The TTSB exhibits oscillations ranging from a minimum of 10 s at high operating temperatures and in the absence of interconnection resistance to a maximum of over 1200 s upon insertion

Table 6
Multivariate linear regression model equations with 3 most influencing β coefficients underlined (1st, 2nd, 3rd) and resulting R^2 values.

Multivariate linear regression models

$$\hat{y} = \beta_0 + \beta_1 \text{Chem}_{NCA} + \beta_2 \text{Chem}_{MIX} + \beta_3 \text{Ageing} + \beta_4 R_{Int} + \beta_5 T_{amb} + \beta_6 R_{Int}^2 + \beta_7 T_{amb}^2 + \beta_8 \text{Chem}_{NCA} \cdot \text{Ageing} + \beta_9 \text{Chem}_{MIX} \cdot \text{Ageing} + \beta_{10} \text{Chem}_{NCA} \cdot T_{amb} + \beta_{11} \text{Chem}_{MIX} \cdot T_{amb} + \beta_{12} \text{Ageing} \cdot R_{Int} + \beta_{13} \text{Ageing} \cdot T_{amb} + \beta_{14} R_{Int} \cdot T_{amb} \tag{17}$$

Response variable	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7
$\%\sigma I_{\text{Start}}$ [−]	0.153	−0.002	0.044	0.014	0.098	0.020	–	–
$\%\sigma I_{\text{Mid}}$ [−]	0.041	−0.014	0.021	0.005	0.020	0.003	0.005	–
$\%\sigma I_{\text{End}}$ [−]	0.151	−0.003	0.032	0.009	0.098	0.025	–	–
ΔSoC_{Max} [%]	9.430	−0.795	2.600	0.659	6.764	0.807	1.093	–
ΔSoC_{End} [%]	2.648	−0.939	0.981	0.287	1.402	−0.683	1.384	–
$\Delta T_{\text{Net}}^{Max}$ [°C]	0.147	–	–	0.313	0.477	−0.941	–	0.552
σT_{Mean} [−]	0.384	−0.052	0.089	0.030	0.171	−0.051	–	–
TTSB [−]	318.575	58.834	−42.268	82.158	359.093	−210.082	−135.242	86.642

Response variable	β_8	β_9	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	R^2
$\%\sigma I_{\text{Start}}$ [−]	0.018	−0.010	−0.005	0.020	–	0.005	–	**0.93**
$\%\sigma I_{\text{Mid}}$ [−]	–	–	–	–	–	–	0.003	**0.91**
$\%\sigma I_{\text{End}}$ [−]	0.016	−0.008	−0.005	0.018	–	–	0.015	**0.90**
ΔSoC_{Max} [%]	0.729	−0.439	–	–	–	0.498	–	**0.97**
ΔSoC_{End} [%]	–	–	0.328	−0.651	–	0.551	−0.478	**0.78**
$\Delta T_{\text{Net}}^{Max}$ [°C]	–	–	–	–	–	− 0.247	–	**0.61**
σT_{Mean} [−]	–	–	–	–	–	–	−0.055	**0.67**
TTSB [−]	–	–	−55.121	82.870	43.473	−46.775	−107.180	**0.94**

of an aged cell. The combination of NMC and NCA cells in parallel is confirmed to be possible. Nevertheless, a detrimental effect on current and temperature distribution is noted. This is mainly attributable to the different internal characteristics of the two cell typologies. Despite the discharge capacity difference being in the order of 2%, the internal resistance of the NCA cells is 30% lower than the NMC ones. This results in added heterogeneity when cells are connected in parallel. Similarly, the inclusion of an aged cell negatively impacts the balance between the cells. The lower discharge capacity and higher internal resistance of the inserted cell increase load and temperature differences. The influence on the TTSB is instead minimal. The linearity or non-linearity of the input–output relationship is maintained even when different chemistries and an aged cell are inserted in the string.

Although the type of graph presented in Fig. 6 assists in qualitatively interpreting the obtained results and provides information on the potential impacts of control variables, a quantitative method is necessary. For this purpose, the fundamental equations underlying these graphs are analysed through coefficient normalisation. It is then possible to directly compare the impact that the variation of each individual control variable has on the response variables. By applying the Akaike method, the number of control variables can be restricted to only those that are statistically significant. Although not within the scope of this article, equations can be utilised for future optimisation studies. The underlining system included in Table 6 help visualising, for each response variable, the three most relevant control variables. The control variables under consideration are multiplied by 14 coefficients, to which the intercept is added. These 14 coefficients are composed of linear, quadratic, and interaction factors. The chemistry features are divided into two distinct categories, namely $Chem_{NCA}$ and $Chem_{Mix}$. This is necessary because, given the three levels in the DoE (NCA, NMC, NCA-NMC mix), the MLR model treats this scenario as two binary pairs. It is immediately evident that interconnection resistance is the most impactful factor for the examined response variables. Both the initial and final current distribution are significantly dependent on the chemistry and its combinations. In the case of σI_{mean}, chemistry once again plays a significant role. Subsequently, it is further explored as to how chemistry and ageing can be attributed to the internal characteristics of cells. This makes the trends of the models more comprehensible by replacing categorical variables with numerical ones. Regarding the maximum and final difference in cells' SoC, the interconnection resistance and its square, as well as the inclusion of different chemistries, play a primary role. The performance of the obtained models is generally good, with five out of the eight response variables presenting a R^2 of 0.9 in an 80/20 train/test split. The predictive models for temperature have yielded less satisfactory results. This is mainly attributable to the measurement uncertainty of the temperature itself, with the thermocouples implemented being limited in tracking the minimal temperature variations under the considered 0.75C discharge conditions. However, it is possible to infer from the equations that the interconnection resistance, operating temperature and its square affect both the net surface temperature increase and its gradient between the cells. Finally, the TTSB can be primarily attributed to interconnection resistance, its square and operating temperature.

As briefly mentioned in the preceding paragraph, it is challenging to linearise categorical variables. An analysis is conducted to investigate the potential association between categorical variables of chemistry and ageing with numerical variables of internal resistance and nominal capacity. This is physically justified by the evolution of cells over time. As SoH deteriorates, internal resistance inevitably increases and the amount of exchangeable energy decreases. A linear model with internal resistance and capacitance has been proposed as an alternative to ageing and chemistry variables for control purposes. The correlations observed between ΔR_0, ΔC_n, chemistry, and ageing are consistent with the cells physics. To prevent potential correlations between ΔR_0 and ΔC_n that may affect the models, it is ensured that the statistical value of Variance Inflation Factor (VIF) is always less than 5 [97]. To determine the VIF value, a regression analysis is conducted on the remaining variables included in the model to predict the variable of interest. A high VIF indicates a high coefficient of determination (R^2) for the given correlation. In essence, if a change in the independent variable has an impact, the effect can already be captured by the other variables included in the model, as their linear combination can approximately account for it. It is important to limit correlations among the control variables concerning the PDP plots that are presented at the end of this section. By replacing categorical variables with numerical variables, it is possible to obtain a greater level of detail regarding the variation in the space of control variables, as they are no longer limited to two or three levels.

Figs. 7, 8, 9 present contour plots resulting from the linear model for TTSB. The x and y axes are replaced by the new control variables ΔR_0 and ΔC_n. The columns indicate the three levels of interconnection resistance, while the rows represent the test temperatures. Contour plots are constrained by the observations obtained in the tests. Thus, extrapolations beyond the tested conditions are not accounted for to

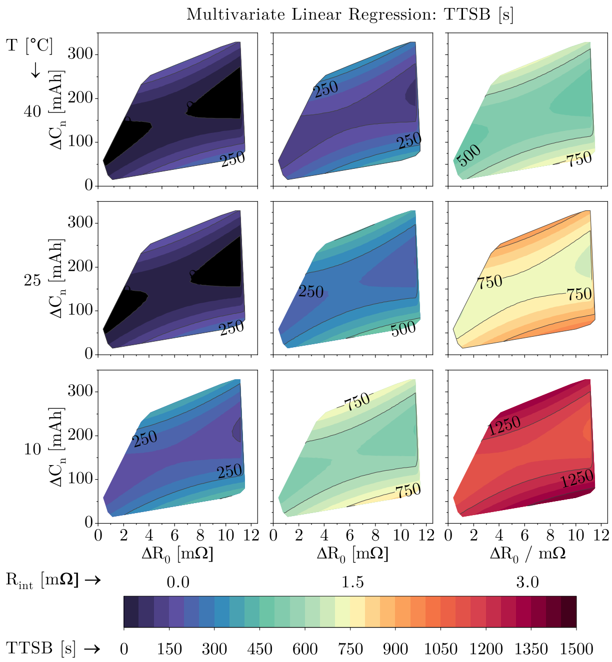

Fig. 7. Contour plots derived from the multivariate linear regression model simulated space for TTSB.

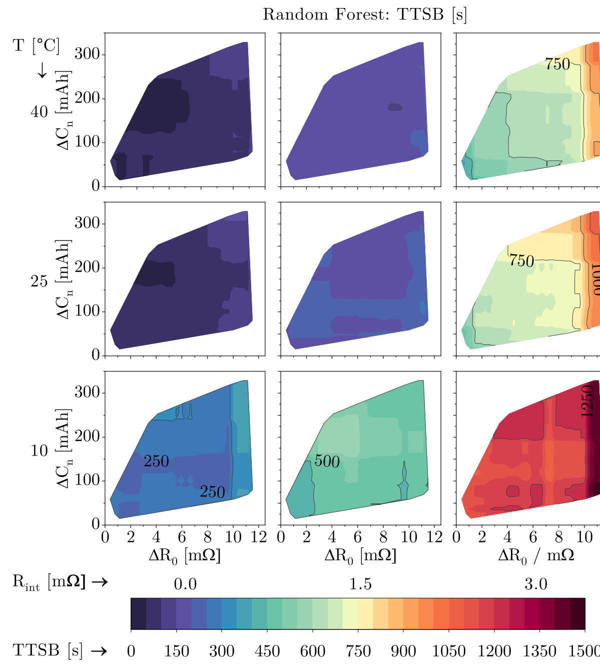

Fig. 9. Contour plots derived from the Random Forest model simulated space for TTSB.

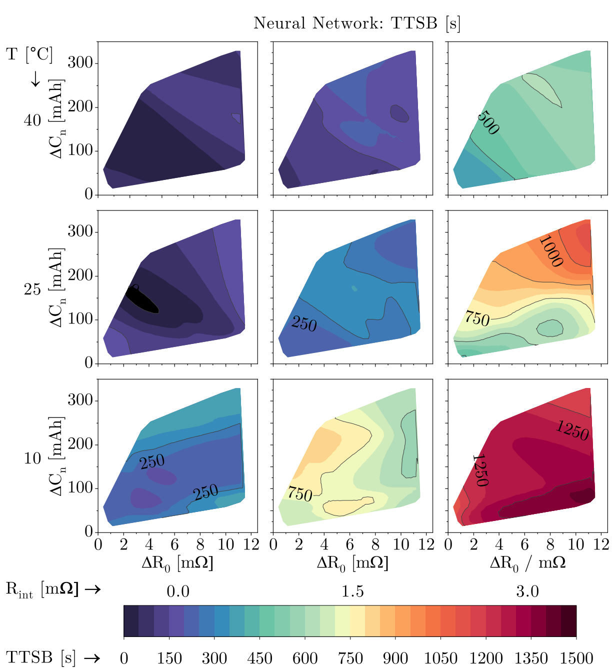

Fig. 8. Contour plots derived from the Neural Network model simulated space for TTSB.

Table 7
Evaluated 5-fold cross-validation models performance mean values and standard deviations.

Response variable	Model mean (Std) R^2		
	MLR	NN	RF
σI_{Start}	0.91(0.03)	0.90(0.13)	0.90(0.04)
σI_{Mid}	0.73(0.19)	0.74(0.13)	0.89(0.05)
σI_{End}	0.86(0.08)	0.90(0.04)	0.85(0.16)
ΔSoC_{Max}	0.91(0.04)	0.87(0.14)	0.93(0.02)
ΔSoC_{End}	0.47(0.24)	0.59(0.73)	0.53(0.11)
ΔT^{Max}_{Net}	0.43(0.25)	0.70(0.60)	0.46(0.27)
σT_{Mean}	0.49(0.23)	0.56(0.28)	0.55(0.43)
$TTSB$	0.73(0.19)	0.65(0.20)	0.66(0.13)

enhance the level of reliability of the presented information. The same control variables are employed for training and validation of the two machine learning models. These are MLP neural networks for Fig. 8 and Random Forest for Fig. 9. The distinction between the conventional linear models previously discussed and ML models lies in their interpretability. It is evident from Figs. 8 and 9 that the two ML models are more influenced by the dataset used. This is evident by the islands that form in the NN and by the square contours resulting from the Random Forest. If the purpose of this article had solely been to predict the performance of parallel cells, it would not have made a difference. The predictive performances of the three models are reported in Table 7, which do not significantly differ from each other. An algorithm of K-fold validation is applied. This presents a more stringent scenario than a classic train/test 80–20 split. Table 7 presents the R^2 results as the average of 50 simulations for more reliable conclusions, with the corresponding standard deviation reported in brackets. The standard deviation is significant as it indicates the stability of the outcome when varying the underlying data sets. It emerges that for those variable responses such as temperature and final SoC variation, not only is the absolute predictive performance limited, but also its repeatability. In line with the objective of shedding light on the most significant factors leading to the imbalanced performance of parallel cells, it is necessary to conduct a more in-depth investigation of the features under consideration.

As machine learning models lack equations that are directly interpretable, methods of explainable machine learning are increasingly gaining interest within the scientific community. Even in the case of this study, where only four control variables are considered, the effectiveness of these methods can be observed. The first applied method is the SHAP technique. The SHAP analyses of the eight response variables conducted for the random forest model are presented in Fig. 10. Similar results are obtained using the NN model, which are not reported for the sake of brevity. The SHAP approach enables the creation of a ranking of the most important features and estimation of the impact

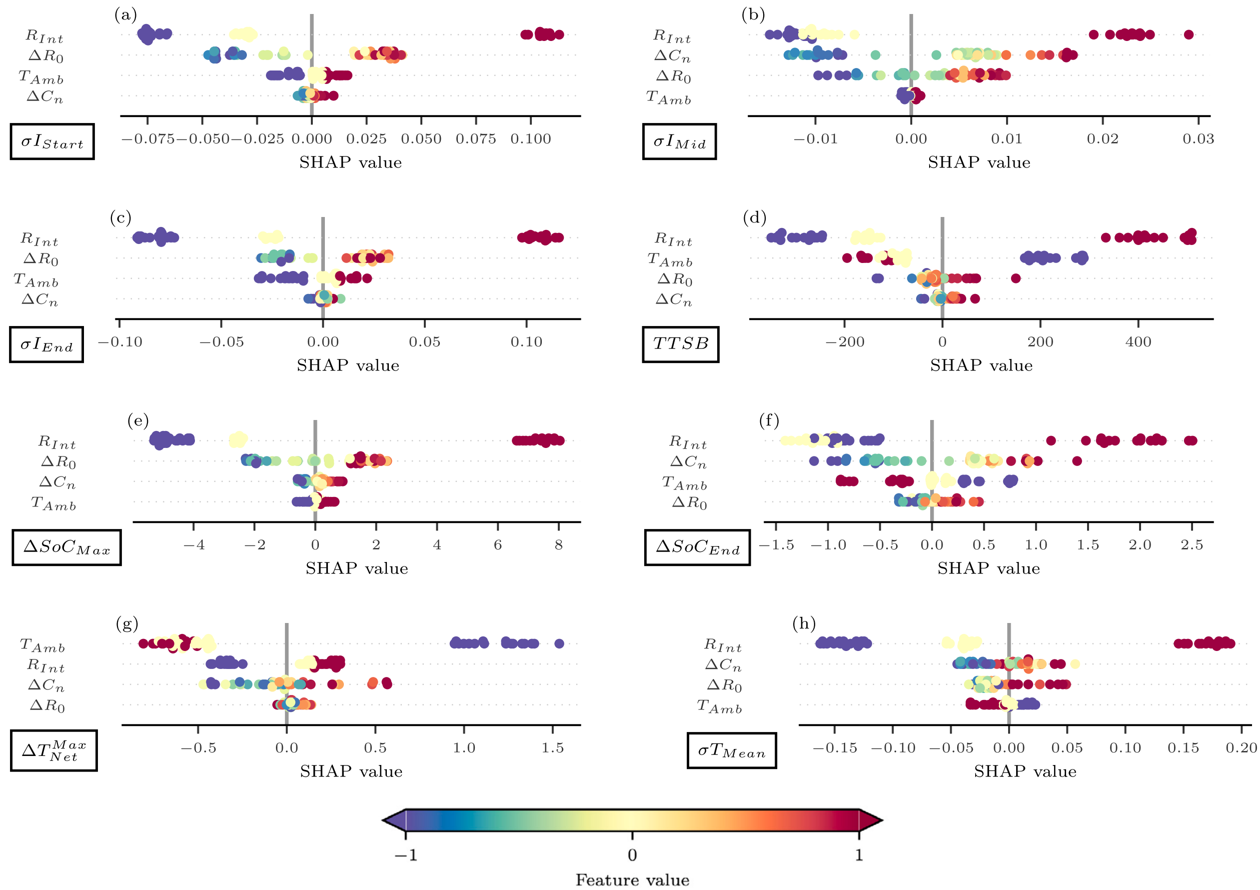

Fig. 10. Random Forest model simulated Shap values for the eight considered response variables (a–h). On the left side of individual graphs the resulting feature importance ranking is reported, while the circles' colour shades indicate their absolute value. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of control variables relative to the mean. The SHAP value is expressed in the same unit of measurement as the response variable. Similarly to linear models, in this case, the interconnection resistance is found to be the primary factor influencing parallel strings performance, except for ΔT_{Net}^{MAx}. Fig. 10(g) illustrates how the increase in cell temperature is inversely affected by the operating temperature. This can be attributed to the fact that the lower the operating temperature, the higher the increase during discharge due to increased ohmic losses. The internal resistance of the cell is known to be a function of the operating temperature. The remaining three positions in the feature ranking vary depending on the considered response. It is noteworthy that in Fig. 10(b), the σI_{Mid} is strongly influenced by the ΔC_n. It can be inferred that the difference in nominal capacity between cells affects the current distribution during the central phase of discharge. The variation of internal resistance at the beginning and end of the discharge has a greater effect, as depicted in Fig. 10(a) and (c). The operating temperature also plays a significant role in the TTSB depicted in Fig. 10(h). The lower the operating temperature, the longer the self-balance currents persist after discharge. The interpretation of the difference in SoC in Fig. 10(e) and (f) is more complicated. Given the model's greater stability in terms of ΔSoC_{Max} performance predictions compared to ΔSoC_{End}, it can be assumed with greater confidence that ΔR_0 has an impact on the peak difference in SoC. In contrast, Fig. 10(f) does not provide a clearly defined ordering of the attributes. Both the control parameter variations that uniformly contribute to the response as well as the model's inability to predict the response with certainty might have an influence. For σT_{Mean} in Fig. 10(h) it can be inferred that interconnection resistance plays a primary role, but it is more difficult to isolate the impact that results from the other control variables.

Despite being effective in providing a ranking of controllable variables, SHAP analysis does not provide information on how their impact is distributed spatially. The PDP and ICE graphics are proposed to fill this gap and further raise the level of interpretability of the models. These two approaches are complementary. The PDP graphics are created by executing the ICE graphic's average. One drawback of PDP plots is that they only display the median of marginal effects. A straight line results if half of the data have a positive relationship to the response variable and half do not. The ICE plots inform on the potential heterogeneous effects and are included to overcome this restriction. The ICE and PDP plots are shown in Fig. 11 for both the NN model and the Random Forest. Five of the eight response variables are reported. The distribution of initial and final current, maximum and final SoC variation, ΔT_{Net}^{Max} and σT_{Mean} are two-by-two correlated. Only the trends of the remaining five responses are being reported for simplicity. The angular coefficient of curves provides information similar to that provided by the linear model's normalised coefficients. As a result, it emerges that the interconnection resistance, following earlier analyses, is the most significant since it results in a greater angular coefficient. Every response variable increases as a result of its rise. The transformation of the previously shown variables is what causes the larger granularity of the ΔR_0 and ΔC_n. The discussion is different concerning temperature and interconnection resistance, both of which have three levels per the DoE configuration. The increase in the number of observations for the ΔR_0 and the ΔC_n has an impact on the ICE graphics, which change depending on the value of the features. The PDP graphics are naturally less susceptible to these variations and may provide better trend information. The two ML models draw conclusions dealing similarly with the control variables. As predicted in the SHAP analysis, while addressing σI_{Mid}, the ΔC_n plays a primary role

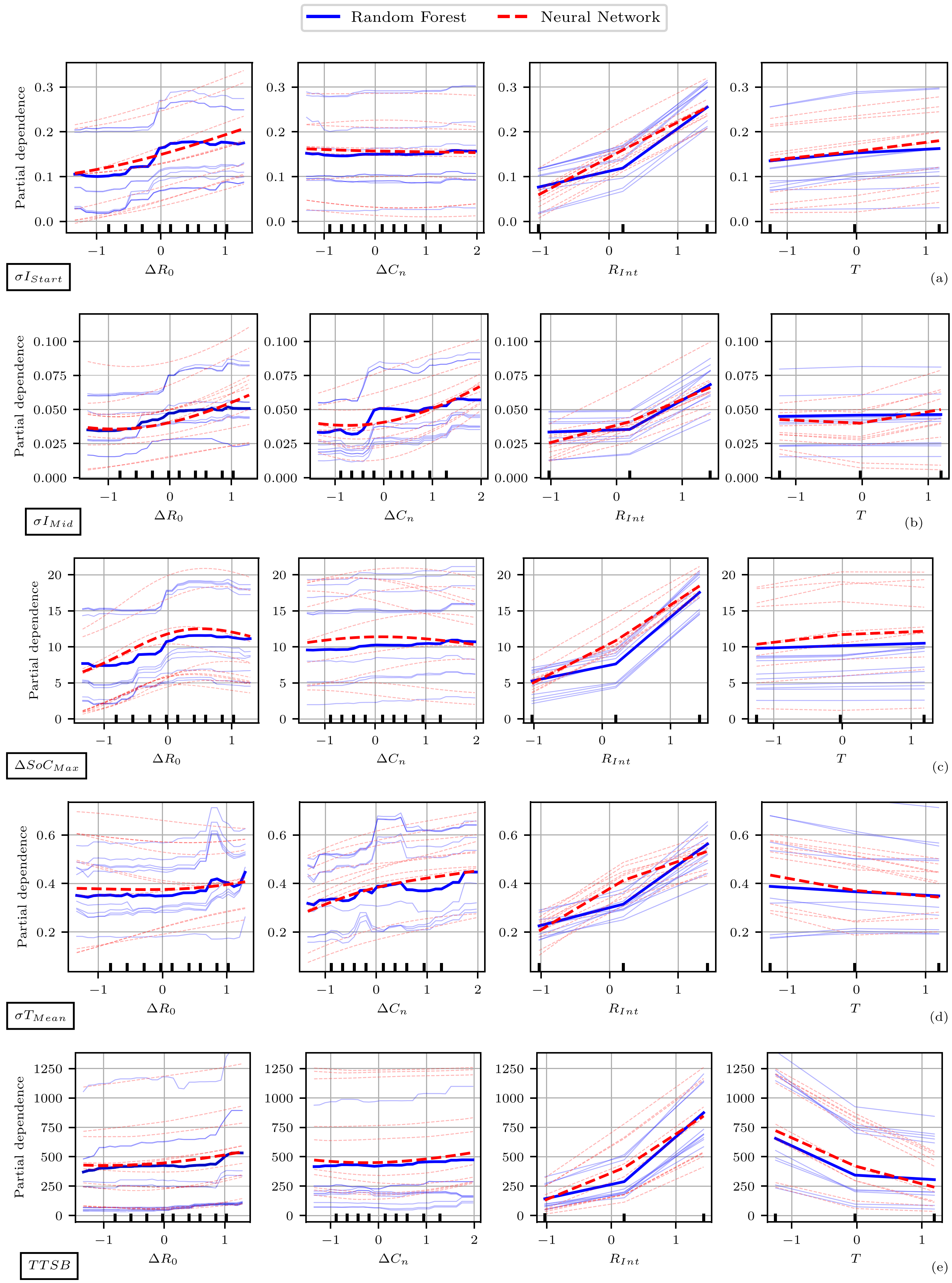

Fig. 11. Random Forest (red) and Neural Network (blue) models PDP and ICE plots for five out of the eight response variables (a–e). Each column reports the PDP and ICE plots resulting from one individual feature. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

together with the interconnection resistance, as shown in Fig. 11(b). When it comes to the σI_{Start} in Fig. 11(a), however, the ΔC_n and the operating temperature have secondary roles, with the interconnection resistance and the ΔR_0 playing the most significant ones. In the case of NN, there is a strong linear relationship between ΔR_0 and σI_{Start}. In the case of Random Forest, there is a tendency to see increases only in values of ΔR_0 close to the median. This pattern is repeated for the ΔSoC_{Max} in Fig. 11(c). The x-axis displays normalised values when feature scaling is used before applying the models. This makes a direct comparison between the controllable variables possible. This kind of analysis is useful when trying to determine not only which variables are most significant but also if their effects change as a function of their values. It is difficult to fully understand the partial dependency of σT_{Mean} on ΔR_0 and ΔC_n in Fig. 11(d) in the case of Random Forest. Oscillations over the whole space might indicate a clear susceptibility to the underlying dataset. Finally, the TTSB is directly correlated with interconnection resistance in Fig. 11(e) and inversely correlated with exercise temperature.

It is possible to obtain a graphic for each controllable variable. PDP and ICE plots provide information on the impact of each feature included in the analysis. However, there is no information on how these entities interact. A limitation of 1D PDPs is the requirement for control variables to be independent to be effective. To incorporate the interaction between control variables, it is possible to utilise two-dimensional PDP graphs. As depicted in Fig. 12, these graphs bear resemblance to the contour plots presented in Fig. 6. Nevertheless, they differ in that they do not directly present the response variables, but rather the marginal effect that the two features have on the predicted response. The x and y axes represent the control variables under investigation, while the colours represent the marginal variation of the response. Not only the colour but also the curvatures provide insight into the relationships among the features. Similarly to one-dimensional PDPs, five response variables are reported for two-dimensional PDPs. The five responses include eight plots each. The first row of the eight images is derived from Random Forest models, while the second one is generated from NN models. The first three columns show the relationship between the most significant control variable derived from the SHAP analysis and the second, third and fourth one, respectively. The last column depicts the interaction between the second and third-ranked control variables. The x and y axes are constrained to values between -1 and 1 as the features are normalised. Both models are capable of creating nonlinear relationships between the control variables. This adds value in particularly complicated situations, like the one being discussed. The Random Forest model tends to have a squared structure, which suggests that they may be more prone to overfitting. The NNs provide a more understandable topology. Their curves indicate first-order and quadratic interactions between the control variables. For instance, in the case of σI_{Mid} in Fig. 12(b), the relationship between interconnection resistance and ΔC_n exhibits a strong curvature. This implies that the higher the variation in nominal capacity, the lower the impact of interconnection resistance. At low levels of nominal capacity variation, the impact of interconnection resistance is predominant in vice versa. Similarly, the same reasoning can be applied to ΔR_0 when compared to interconnection resistance. The behaviour of the Random Forest model in this case is more challenging to interpret. The surfaces and demarcation lines exhibit rapid and non-uniform variations. These results appear to be squared again, thereby limiting the interpretability of the relationships among the control variables. The combined behaviour of the control variables concerning σI_{Start} in Fig. 12(a) is more linear. An exception arises in the relationship between interconnection resistance and ΔR_0 in the case of Random Forest, which exhibits singularity towards the mean value. Fig. 12(c) displays a curvature in the relationship between interconnection resistance, ΔR_0, and ΔC_n in ΔSoC_{Max}, highlighting a non-linear contribution from the combination of control variables. The curvature between temperature and interconnection resistance is not immediately evident, with the latter being predominant, especially at high levels. A minimum is observed towards the average values for both ΔR_0 and ΔC_n in relation to interconnection resistance. This peculiarity is also present in random forests, which however lose the curvature of NNs and exhibit a strong non-linearity in the shape of an "S". This implies that for values above the average of ΔR_0, the ΔSoC_{Max} is higher than for values below the average, given the same interconnection resistance. The 2D PDP in Fig. 12(d) shows strong oscillations for the Random Forest model employed in predicting σT_{Mean}, thereby constraining its interpretability. In such cases, it is recommended to refer to the NN, whose curvatures are more easily identifiable. The relationships between the control variables for the TTSB in Fig. 12(e) are more linear. The interaction between interconnection resistance and temperature appears to be balanced, confirming previous analyses. The impact of ΔR_0 and ΔC_n on the response is limited. The worst condition occurs under low temperatures and high levels of interconnection resistance. Conversely, the minimum duration for which self-balance currents persists after discharge is at high temperatures with minimum levels of interconnection resistance.

Via these graphs, it is possible to understand whether the ML model utilises linear, monotonic, or non-linear relationships between the control and response variables. In the investigated scenario, the relationships appear to be minimally non-linear. In such cases, a linear model may already be sufficient in predicting both the responses and the importance of the considered features. However, the value of the proposed methodology lies in its independence from the considered data. This assumes additionally greater value in applications where a strong non-linearity is present. Besides, PDP and ICE plots inform about whether the number of observations is adequate or not. In the case of linear relationships, a reduction of the levels of the DoE is possible, with advantages in terms of time and resources. When non-linear relationships emerge, a higher number of observations might instead be required. The management of the experimental space extension is critical in scenarios where the investigated system is costly or the resources available are limited, as in the case of single cells to modules/pack scalability studies. The main advantage of adding ML models to the tools available in feature importance analysis resides in their capability to capture highly non-linear relationships. The application of XML techniques reduces the distance between ML and traditional methods' interpretability, adding a valuable tool towards battery research advancements.

4.1. Limitations and further work

Some limitations emerge from this study. The PDP plots rely on the independence of features. In reality, it is not always possible to ensure this behaviour. Alternative solutions are the Principal Component Analysis (PCA). Nevertheless, the feature transformation would not allow a straight identification of the variables included in the DoE. Despite the effort of limiting and tracking any experimental setup influence, parallel-connected cells are particularly sensitive to the connections. Industry-level busbars do not present bolted connections, with soldering being the selected solution. The results of this study are still relevant in these applications, but the impact of the interconnection resistance is expected to be lower due to its lower absolute value. The methodology is data agnostic in nature but ML models hyperparameters tuning is not. Adequate tuning needs to be considered when applying the methodology to other scenarios.

This article shed light on single cells and module-level features influencing the imbalanced performance of strings. Future work is still required to extend the investigated space to further input and output variables. These could include different architectures (Z, mixed), chemistries (LFP, LMO), form factors (18650, 4650, pouch) and loads (real-world cycles). While the presented study focused on a fixed C-rate during the experimental design, it is acknowledged that variations in C-rate can significantly influence cells' behaviour. Future investigations could benefit from incorporating the C-rate as an additional

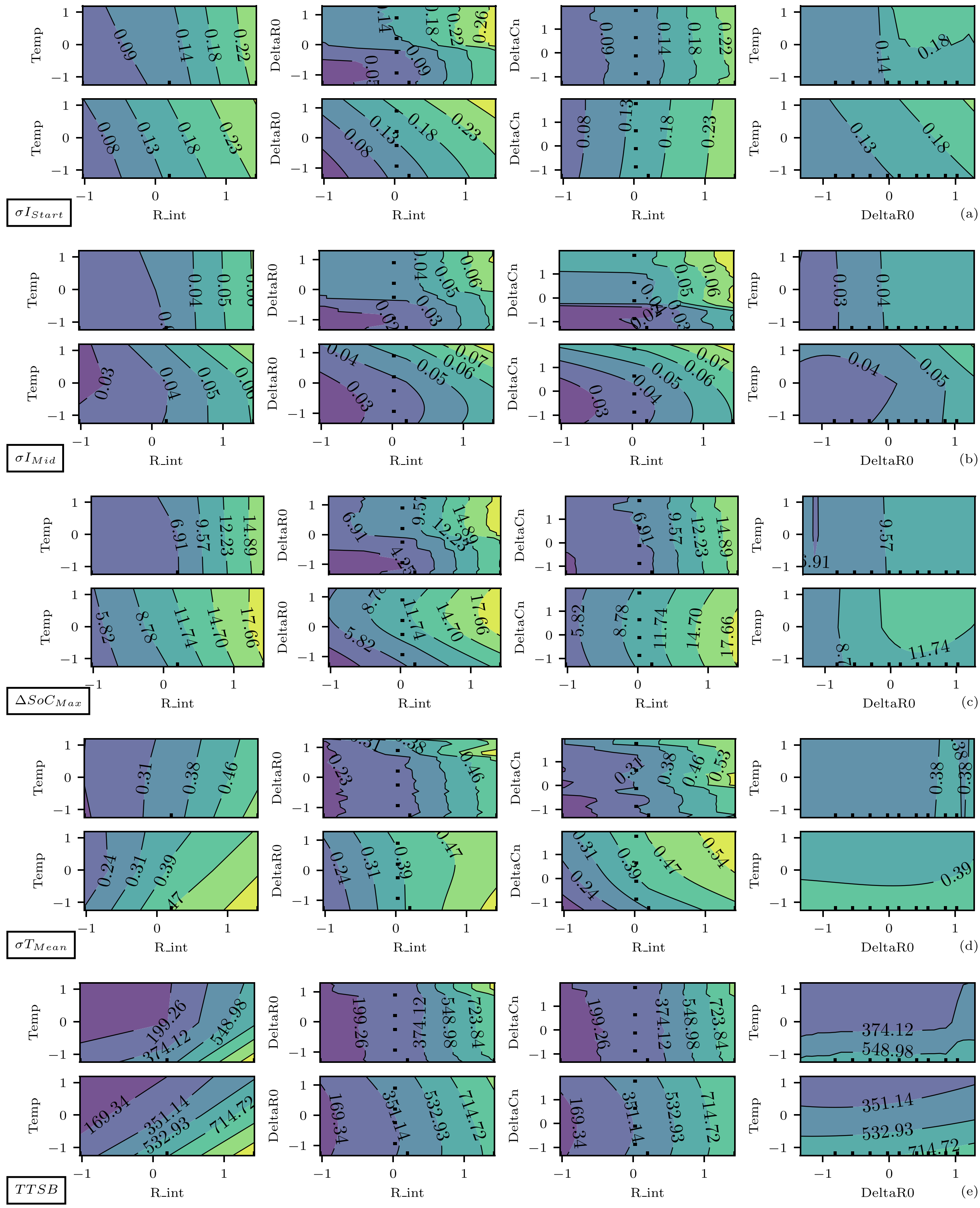

Fig. 12. Random Forest (top) and Neural Network (bottom) models 2-dimensional interaction PDP plots for five out of the eight response variables (a–e). Each column includes the resulting plot from the interaction of 2 features in decreasing importance order derived from the SHAP analysis (1st-2nd, 1st-3rd, 1st-4th, 2nd-3rd). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

independent variable to enhance the generalisability of the findings across various operating conditions. The heterogeneities included in this paper are of a performance nature. The implications that the protraction of imbalances can have on the cells' ageing evolution is an important perspective still missing a unified understanding in the literature. Coherently, their influence on SoH distribution evolution and ageing rate is subject to future investigation. The prediction ability emerging from the evaluated models can be leveraged to reduce the number of experiments in future studies. The performed full-factorial DoE includes an I-optimal design in its first 30 runs. The comparison of the prediction ability with a subset of the data can allow test number reduction in costly and resource-intensive conditions. When scaling from single cells to packs, it is not possible to maintain the same testing approach and extent due to the capital costs involved. The proposed solution can reduce the investigation space and future work is required to demonstrate the potential, especially in resource-intensive campaigns such as ageing ones. Last, the set of equations generated by the MLR model can be leveraged for optimisation purposes. Busbar and cooling design, manufacturing quality control and cells' second-life suitability selection are fields with high applicability potential for future studies. Besides, the collected experimental data will be play a twofold role in future activities. Namely, enabling the development of data-driven methodologies for parallel string performance prediction and providing validation datasets for module-level modelling formulations.

5. Conclusions

In this paper the impact of cell and module-level properties on the performance of parallel connected strings is experimentally investigated. Increasing the level of understanding of the most important features influencing the heterogeneity propagation inside the modules/packs is crucial to limiting the phenomenon. A methodology combining well-established non-invasive single-cell characterisation tests with data-driven modelling tools is proposed.

Two batches of twenty new NMC and NCA cells are first characterised to identify out-of-manufacture internal resistance and capacity distributions. The NCA cells' discharge capacity is normally distributed around the mean. Some NMC cells are categorised as outliers, otherwise causing a deviation in the distribution of the fresh batch and subsequently to the module-level tests. The standard deviation across cells is found to be a function of the SoC. Not only the absolute value of the internal resistance but also its standard deviation increases at low SoC. A larger dispersion of properties worsens the performance homogeneity of parallel strings. Coherently, investigating the parallel strings' heterogeneity as a function of SoC could be of relevant interest and left for future investigation.

A 54-test condition full-factorial DoE campaign is conducted on four ladder-parallel connected modules to consider all components independently and thereby discern their influence. The experiments inform about how the cells' current, State of Charge and temperature distributions and time to self-balance under 0.75C constant current discharge loads are affected by interconnection resistance, operating temperature, different chemistry combination and ageing. The most influential elements in modules' unbalanced performance are identified. The experimental results are used to create a multivariate linear regression model that links the most important control factors to the chosen response variables. Explainable machine learning techniques are compared with conventional linear regression analysis. Multi-layer perceptron neural networks and random forest machine learning models are trained and tested to get alternative information to multiple linear regression, outlining their benefits and drawbacks. The key findings can be summarised as follows:

I. **The DoE statistical analysis results underline that combining NMC and NCA cells in parallel is possible.** Nevertheless, an added level of attention needs to be put into the management of the system due to generally increasing performance inhomogeneities. The current deviation across the cells doubles when compared to the single chemistry condition. Temperature gradients increase, while the TTSB is not particularly influenced by their combination.
II. **Introducing an aged cell in the string worsens the homogeneity of the cells' heterogeneous performance.** The operating temperature impact on the imbalances increases, suggesting more focus on the thermal management might be required in this case. Although heterogeneities get worse, the trends are not strongly impacted. Linear and non-linear relationships between control and response variables are maintained with and without an aged cell insertion in the string.
III. **The combination of novel XML techniques and established statistical analysis allows unravelling features contribution to parallel cells performance.** According to the results, the interconnection resistance is the most relevant contributor to strings' heterogeneous performance. In the first and middle phases of the discharge, the distributions of internal resistance and capacity impact the load imbalance across the cells, respectively. Increasing the operating temperature negatively influences the temperature gradient in the string. To mitigate self-balancing currents after the discharge, adequate busbar design and thermal management are fundamental.
IV. **In all but temperature-sensed responses and TTSB, linear models are a sufficient solution for their simplicity and optimisation potential.** The relationship between control and response variables appears generally linear, with the exceptions being the aforementioned responses. Nevertheless, the application of XML techniques reduces the interpretation complexity of ML models and underlines their potential, especially for cases when non-linearities are dominant. PDP and ICE plots show that in linear regions, few levels are sufficient. Conversely, a finer resolution is necessary to represent non-linear dependencies as TTSB.

It has been shown that in the evaluated cases traditional linear regression and machine learning models show similar prediction performance. All the models have a R^2 value above 0.9, with the ΔSoC_{End} and the temperature-related signals being exceptions. The range of R^2 is from 0.74 up to 0.98 for the remaining responses. Despite having a good performance, machine learning models confirm their interpretability limitations. Islands-type and squared surfaces emerge for neural network and random forest responses, respectively. Apart from being a symptom of dataset dependency, it is difficult to directly interpret how the models are dealing with the control variables in these areas. Therefore, XML techniques showcase their potential is overcoming this limitation. The three techniques have their peculiarities, with SHAP being strong in ranking the features, 1D PDP and ICE in explaining the spatial distribution of control-response relationships and 2D PDP in unveiling eventual interactions.

CRediT authorship contribution statement

Gabriele Piombo: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Simone Fasolato:** Data curation, Formal analysis, Investigation, Methodology, Software, Visualization. **Robert Heymer:** Formal analysis, Methodology, Visualization. **Marc Hidalgo:** Methodology. **Mona Faraji Niri:** Methodology, Supervision, Writing – review & editing. **Simona Onori:** Funding acquisition, Resources, Supervision, Writing – review & editing. **James Marco:** Funding acquisition, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

A Data in Brief co-submission will be performed.

Acknowledgements

The research outlined in this paper is supported by Altair Engineering (UK) Ltd. via the Engineering Centre for Doctoral Training in Sustainable Materials and Manufacturing (grant number EP/L016389/1). The authors gratefully acknowledge the support of the Stanford Energy Control Lab for hosting the experimental activities and of Andrew Moore (WMG, University of Warwick) for the sensing setup development.

Appendix A. Notes on the selected ML models

In situations where comprehensibility is crucial Decision Tree (DT) are commonly selected thanks to their adequate levels of model interpretability. However, they have a tendency to overfit, resulting in inadequate generalisation. The Random Forest (RF) algorithm was originally introduced as a means of enhancing the precision of individual DT. This is achieved by aggregation of several trees to enhance the generalisation of the model and thereby resulting in improved performance [86]. To accomplish this, a distinct tree is trained on a specific subset of the training dataset, capturing diverse features of the data distribution, to acquire a combined forecast. This methodology yields highly precise models [98]. In contrast to DT, which RF shares fundamental principles with, RF models lack transparency and are therefore challenging to interpret. XML techniques are gaining increasing interest as a solution to aid in the comprehension of models' decisions.

The selection of the Multi-Layer Perceptron (MLP) as NN method for regression is attributed to its high precision and ability to effectively manage complex and non-linear datasets, despite being the most basic and low computational intensive NN approach with fully connected layers. MLP NN, from a technical perspective, is constructed by arranging layers of nodes that link the input features to the target variable sequentially. It can be observed that every individual node situated in an intermediate layer undertakes the task of gathering and consolidating the outputs generated by the preceding layer. Subsequently, the node generates an output of its own by subjecting the consolidated value to an activation function. These aforementioned values are transmitted to the next layer(s) until the final output layer is attained. The back-propagation technique is utilised in the MLP algorithm to optimise the interneuron weights and enhance the precision of the model. In relation to the scope of this study, it is noteworthy that the MLP algorithm carries out feature selection during the training phase. This results in the exclusion of non-contributing variables during model generation, thereby simplifying feature ranking.

While the paper does not delve into determining the ideal model for the dataset at hand, optimising hyperparameters is crucial in enhancing the accuracy of the ML models and consequently, the efficacy of the features associated with the responses. A Random Search (RS) technique is employed to decrease the search space without exhaustively evaluating all feasible hyperparameter combinations. After, a Grid Search (GS) algorithm is employed to identify the most favourable parameters for all the response variables that are incorporated. The optimisation of hyperparameters does not apply to MLR models, whose equations are not tuneable apart from the feature engineering steps explained above. This reduced design freedom, combined with their simplicity and ease of implementation, made MLR the main approach for DoE statistical analysis.

Supplementary information

The experimental data of this study will be published in Data in Brief [96]. Please contact the authors for further information. The details of the hyperparameter optimisation range for the machine learning models are presented below to ensure transparency and replicability of the contents of the study. The laboratory data management and extraction of response variables are conducted in MATLAB 2022b. The Stat-Ease Design Expert 22.0.2 software is employed to develop the multivariate linear model and analyse the importance of the DoE features. The training, validation of ML models, and feature analysis via XML are carried out using Python3.

- Multi-Layer Perceptron Neural Network : Hidden layer size [(50,50,50), (50,100,50), (100,)], Activation = [tanh relu], Solver = [sgd adam], Alpha = [0.0001, 0.1], Learning rate = [constant adaptive]
- Random Forest: Number of estimators: [50 1000], Bootstrap = [True False], Minimum samples split = [2 3], Minimum samples leaf = [1 3], Criterion = [Squared error]

References

[1] S.J. Davis, N.S. Lewis, M. Shaner, S. Aggarwal, D. Arent, I.L. Azevedo, S.M. Benson, T. Bradley, J. Brouwer, Y.M. Chiang, C.T. Clack, A. Cohen, S. Doig, J. Edmonds, P. Fennell, C.B. Field, B. Hannegan, B.M. Hodge, M.I. Hoffert, E. Ingersoll, P. Jaramillo, K.S. Lackner, K.J. Mach, M. Mastrandrea, J. Ogden, P.F. Peterson, D.L. Sanchez, D. Sperling, J. Stagner, J.E. Trancik, C.J. Yang, K. Caldeira, Net-zero emissions energy systems, Science 360 (6396) (2018) http://dx.doi.org/10.1126/science.aas9793.

[2] X. Hu, C. Zou, C. Zhang, Y. Li, Technological developments in batteries: A survey of principal roles, types, and management needs, IEEE Power Energy Mag. 15 (5) (2017) 20–31, http://dx.doi.org/10.1109/MPE.2017.2708812.

[3] T. Weaver, A. Allam, S. Onori, A novel lithium-ion battery pack modeling framework-series-connected case study, in: Proceedings of the American Control Conference, Vol. 2020-July, Institute of Electrical and Electronics Engineers Inc., 2020, pp. 365–372, http://dx.doi.org/10.23919/ACC45564.2020.9147546.

[4] W. Chen, J. Liang, Z. Yang, G. Li, A review of lithium-ion battery for electric vehicle applications and beyond, Energy Procedia 158 (2019) 4363–4368, http://dx.doi.org/10.1016/j.egypro.2019.01.783.

[5] M. Baumann, L. Wildfeuer, S. Rohr, M. Lienkamp, Parameter variations within li-ion battery packs – theoretical investigations and experimental quantification, J. Energy Storage 18 (2018) 295–307, http://dx.doi.org/10.1016/j.est.2018.04.031.

[6] Y. Liu, R. Zhang, J. Wang, Y. Wang, Current and future lithium-ion battery manufacturing, iScience 24 (4) (2021) 102332, http://dx.doi.org/10.1016/j.isci.2021.102332.

[7] K. Rumpf, M. Naumann, A. Jossen, Experimental investigation of parametric cell-to-cell variation and correlation based on 1100 commercial lithium-ion cells, J. Energy Storage 14 (2017) 224–243, http://dx.doi.org/10.1016/j.est.2017.09.010.

[8] R. Gogoana, M.B. Pinson, M.Z. Bazant, S.E. Sarma, Internal resistance matching for parallel-connected lithium-ion cells and impacts on battery pack cycle life, J. Power Sources 252 (2014) 8–13, http://dx.doi.org/10.1016/j.jpowsour.2013.11.101.

[9] C. Reiter, L. Wildfeuer, N. Wassiliadis, T. Krahl, J. Dirnecker, M. Lienkamp, A holistic approach for simulation and evaluation of electrical and thermal loads in lithium-ion battery systems, in: 2019 Fourteenth International Conference on Ecological Vehicles and Renewable Energies (EVER), Vol. null, IEEE, 2019, pp. 1–17, http://dx.doi.org/10.1109/EVER.2019.8813640.

[10] W. Shi, X. Hu, C. Jin, J. Jiang, Y. Zhang, T. Yip, Effects of imbalanced currents on large-format LiFePO 4 /graphite batteries systems connected in parallel, J. Power Sources 313 (2016) 198–204, http://dx.doi.org/10.1016/j.jpowsour.2016.02.087.

[11] S. Miyatake, Y. Susuki, T. Hikihara, S. Itoh, K. Tanaka, Discharge characteristics of multicell lithium-ion battery with nonuniform cells, J. Power Sources 241 (2013) 736–743, http://dx.doi.org/10.1016/j.jpowsour.2013.05.179.

[12] H. Zhang, T. Wang, R. Lu, C. Zhu, Y. Zhao, Study on the impedance increase fault of parallel connected batteries based on simscape model simulation, in: 2015 IEEE Vehicle Power and Propulsion Conference, VPPC 2015 - Proceedings, Institute of Electrical and Electronics Engineers Inc., 2015, http://dx.doi.org/10.1109/VPPC.2015.7353027.

[13] T. Bruen, J. Marco, Modelling and experimental evaluation of parallel connected lithium ion cells for an electric vehicle battery system, J. Power Sources 310 (2016) 91–101, http://dx.doi.org/10.1016/j.jpowsour.2016.01.001.

[14] C. Pastor-Fernández, T. Bruen, W. Widanage, M. Gama-Valdez, J. Marco, A study of cell-to-cell interactions and degradation in parallel strings: Implications for the battery management system, J. Power Sources 329 (2016) 574–585, http://dx.doi.org/10.1016/j.jpowsour.2016.07.121.

[15] A. Cordoba-Arenas, S. Onori, G. Rizzoni, A control-oriented lithium-ion battery pack model for plug-in hybrid electric vehicle cycle-life studies and system design with consideration of health management, J. Power Sources 279 (2015) 791–808, http://dx.doi.org/10.1016/j.jpowsour.2014.12.048.

[16] A. Fill, T. Schmidt, T. Mader, R. Llorente, A. Avdyli, K.P. Birke, New semi-analytical model approach of the current distribution within parallel-connected lithium-ion cells, J. Energy Storage 40 (2021) 102653, http://dx.doi.org/10.1016/j.est.2021.102653.

[17] L. Chang, C. Wang, C. Zhang, L. Xiao, N. Cui, H. Li, J. Qiu, A novel fast capacity estimation method based on current curves of parallel-connected cells for retired lithium-ion batteries in second-use applications, J. Power Sources 459 (2020) 227901, http://dx.doi.org/10.1016/j.jpowsour.2020.227901.

[18] P. Jocher, M. Steinhardt, S. Ludwig, M. Schindler, J. Martin, A. Jossen, A novel measurement technique for parallel-connected lithium-ion cells with controllable interconnection resistance, J. Power Sources 503 (2021) 230030, http://dx.doi.org/10.1016/j.jpowsour.2021.230030.

[19] A. Fill, T. Mader, T. Schmidt, R. Llorente, K.P. Birke, Measuring test bench with adjustable thermal connection of cells to their neighbors and a new model approach for parallel-connected cells, Batteries 6 (1) (2019) 2, http://dx.doi.org/10.3390/batteries6010002.

[20] C. Li, N. Cui, L. Chang, Z. Cui, H. Yuan, C. Zhang, Effect of parallel connection topology on air-cooled lithium-ion battery module: Inconsistency analysis and comprehensive evaluation, Appl. Energy 313 (2022) 118758, http://dx.doi.org/10.1016/j.apenergy.2022.118758.

[21] A. Fill, S. Koch, K.P. Birke, Analytical model of the current distribution of parallel-connected battery cells and strings, J. Energy Storage 23 (2019) 37–43, http://dx.doi.org/10.1016/j.est.2019.02.031.

[22] W. Diao, M. Pecht, T. Liu, Management of imbalances in parallel-connected lithium-ion battery packs, J. Energy Storage 24 (2019) 100781, http://dx.doi.org/10.1016/j.est.2019.100781.

[23] C. Luan, C. Ma, C. Wang, L. Chang, L. Xiao, Z. Yu, H. Li, Influence of the connection topology on the performance of lithium-ion battery pack under cell-to-cell parameters variations, J. Energy Storage 41 (2021) 102896, http://dx.doi.org/10.1016/j.est.2021.102896.

[24] T. Grün, K. Stella, O. Wollersheim, Influence of circuit design on load distribution and performance of parallel-connected lithium ion cells for photovoltaic home storage systems, J. Energy Storage 17 (2018) 367–382, http://dx.doi.org/10.1016/j.est.2018.03.010.

[25] L. Chang, C. Ma, Y. Zhang, H. Li, L. Xiao, Experimental assessment of the discharge characteristics of multi-type retired lithium-ion batteries in parallel for echelon utilization, J. Energy Storage 55 (2022) 105539, http://dx.doi.org/10.1016/j.est.2022.105539.

[26] Y. Tian, Z. Huang, X. Li, J. Tian, Parallel-connected battery module modeling based on physical characteristics in multiple domains and heterogeneous characteristic analysis, Energy 239 (2022) 122181, http://dx.doi.org/10.1016/j.energy.2021.122181.

[27] N. Yang, X. Zhang, B. Shang, G. Li, Unbalanced discharging and aging due to temperature differences among the cells in a lithium-ion battery pack with parallel combination, J. Power Sources 306 (2016) 733–741, http://dx.doi.org/10.1016/j.jpowsour.2015.12.079.

[28] M. Fleckenstein, O. Bohlen, M.A. Roscher, B. Bäker, Current density and state of charge inhomogeneities in li-ion battery cells with LiFePO4 as cathode material due to temperature gradients, J. Power Sources 196 (10) (2011) 4769–4778, http://dx.doi.org/10.1016/j.jpowsour.2011.01.043.

[29] B. Wu, V. Yufit, M. Marinescu, G.J. Offer, R.F. Martinez-Botas, N.P. Brandon, Coupled thermal–electrochemical modelling of uneven heat generation in lithium-ion battery packs, J. Power Sources 243 (2013) 544–554, http://dx.doi.org/10.1016/j.jpowsour.2013.05.164.

[30] M. Naylor Marlow, J. Chen, B. Wu, Degradation in parallel-connected lithium-ion battery packs under thermal gradients, Communications Engineering 3 (1) (2024) 2, http://dx.doi.org/10.1038/s44172-023-00153-5.

[31] M.P. Klein, J.W. Park, Current distribution measurements in parallel-connected lithium-ion cylindrical cells under non-uniform temperature conditions, J. Electrochem. Soc. 164 (9) (2017) 1893, http://dx.doi.org/10.1149/2.0011709jes.

[32] A. Fill, S. Koch, K.P. Birke, Algorithm for the detection of a single cell contact loss within parallel-connected cells based on continuous resistance ratio estimation, J. Energy Storage 27 (2020) 101049, http://dx.doi.org/10.1016/j.est.2019.101049.

[33] L. Wang, Y. Xu, E. Wang, X. Zhao, S. Qiao, G. Li, H. Sun, Modeling and state of charge estimation of inconsistent parallel lithium-ion battery module, J. Energy Storage 51 (2022) 104565, http://dx.doi.org/10.1016/j.est.2022.104565.

[34] X. Liu, W. Ai, M. Naylor Marlow, Y. Patel, B. Wu, The effect of cell-to-cell variations and thermal gradients on the performance and degradation of lithium-ion battery packs, Appl. Energy 248 (2019) 489–499, http://dx.doi.org/10.1016/j.apenergy.2019.04.108.

[35] Y. Zhang, J. Zheng, S. Lin, F. Bai, W.H. Tanveer, S. Cha, X. Wu, W. Feng, Nonuniform current distribution within parallel-connected batteries, Int. J. Energy Res. 42 (8) (2018) 2835–2844, http://dx.doi.org/10.1002/er.4039.

[36] G.J. Offer, V. Yufit, D.A. Howey, B. Wu, N.P. Brandon, Module design and fault diagnosis in electric vehicle batteries, J. Power Sources 206 (2012) 383–392, http://dx.doi.org/10.1016/j.jpowsour.2012.01.087.

[37] M. Dubarry, A. Devie, B.Y. Liaw, Cell-balancing currents in parallel strings of a battery system, J. Power Sources 321 (2016) 36–46, http://dx.doi.org/10.1016/j.jpowsour.2016.04.125.

[38] C. Campestrini, P. Keil, S.F. Schuster, A. Jossen, Ageing of lithium-ion battery modules with dissipative balancing compared with single-cell ageing, J. Energy Storage 6 (2016) 142–152, http://dx.doi.org/10.1016/j.est.2016.03.004.

[39] M. Schindler, P. Jocher, A. Durdel, A. Jossen, Analyzing the aging behavior of lithium-ion cells connected in parallel considering varying charging profiles and initial cell-to-cell variations, J. Electrochem. Soc. 168 (9) (2021) 090524, http://dx.doi.org/10.1149/1945-7111/ac2089.

[40] X. Wang, Z. Wang, L. Wang, Z. Wang, H. Guo, Dependency analysis and degradation process-dependent modeling of lithium-ion battery packs, J. Power Sources 414 (2019) 318–326, http://dx.doi.org/10.1016/j.jpowsour.2019.01.021.

[41] F. An, J. Huang, C. Wang, Z. Li, J. Zhang, S. Wang, P. Li, Cell sorting for parallel lithium-ion battery systems: Evaluation based on an electric circuit model, J. Energy Storage 6 (2016) 195–203, http://dx.doi.org/10.1016/j.est.2016.04.007.

[42] R. Spurrett, C. Thwaite, A. Holland, D. Lizius, G.J. Dudley, Modeling of highly-parallel lithium-ion batteries, in: Space Power, Vol. 502, 2002, p. 685, URL https://adsabs.harvard.edu/pdf/2002esasp.502..685s.

[43] M. Brand, D. Quinger, G. Walder, A. Jossen, M. Lienkamp, Ageing inhomogeneity of long-term used BEV-batteries and their reusability for 2nd-life applications, in: EVS26 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium, 2012, URL https://mediatum.ub.tum.de/doc/1161064/document.pdf.

[44] S.F. Schuster, M.J. Brand, P. Berg, M. Gleissenberger, A. Jossen, Lithium-ion cell-to-cell variation during battery electric vehicle operation, J. Power Sources 297 (2015) 242–251, http://dx.doi.org/10.1016/j.jpowsour.2015.08.001.

[45] X. Gong, R. Xiong, C.C. Mi, Study of the characteristics of battery packs in electric vehicles with parallel-connected lithium-ion battery cells, IEEE Trans. Ind. Appl. 51 (2) (2015) 1872–1879, http://dx.doi.org/10.1109/TIA.2014.2345951.

[46] N. Kakimoto, K. Goto, Capacity-fading model of lithium-ion battery applicable to multicell storage systems, IEEE Trans. Sustain. Energy 7 (1) (2016) 108–117, http://dx.doi.org/10.1109/TSTE.2015.2476476.

[47] P.-L. Huynh, Beitrag zur Bewertung des Gesundheitszustands von Traktionsbatterien in Elektrofahrzeugen, Springer Fachmedien Wiesbaden, 2016, http://dx.doi.org/10.1007/978-3-658-16562-8.

[48] M. Duquesnoy, T. Lombardo, M. Chouchane, E.N. Primo, A.A. Franco, Data-driven assessment of electrode calendering process by combining experimental results, in silico mesostructures generation and machine learning, J. Power Sources 480 (2020) 229103, http://dx.doi.org/10.1016/j.jpowsour.2020.229103.

[49] K. Rumpf, A. Rheinfeld, M. Schindler, J. Keil, T. Schua, A. Jossen, Influence of cell-to-cell variations on the inhomogeneity of lithium-ion battery modules, J. Electrochem. Soc. 165 (11) (2018) A2587–A2607, http://dx.doi.org/10.1149/2.0111811jes.

[50] E. Hosseinzadeh, S. Arias, M. Krishna, D. Worwood, A. Barai, D. Widanalage, J. Marco, Quantifying cell-to-cell variations of a parallel battery module for different pack configurations, Appl. Energy 282 (2021) 115859, http://dx.doi.org/10.1016/j.apenergy.2020.115859.

[51] M. Shahjalal, P.K. Roy, T. Shams, A. Fly, J.I. Chowdhury, M.R. Ahmed, K. Liu, A review on second-life of li-ion batteries: prospects, challenges, and issues, Energy 241 (2022) 122881, http://dx.doi.org/10.1016/j.energy.2021.122881.

[52] K. He, S. Tao, S. Fu, H. Fan, Y. Tao, Y. Wang, Y. Sun, A novel quick screening method for the second usage of parallel-connected lithium-ion cells based on the current distribution, J. Electrochem. Soc. 170 (3) (2023) 030514, http://dx.doi.org/10.1149/1945-7111/ACBF7E.

[53] S. Yang, C. Zhang, J. Jiang, W. Zhang, H. Chen, Y. Jiang, D.U. Sauer, W. Li, Fast screening of lithium-ion batteries for second use with pack-level testing and machine learning, eTransportation 17 (2023) 100255, http://dx.doi.org/10.1016/j.etran.2023.100255.

[54] Z. Li, A. Zuo, Z. Mo, M. Lin, C. Wang, J. Zhang, M.H. Hofmann, A. Jossen, Demonstrating stability within parallel connection as a basis for building large-scale battery systems, Cell Rep. Phys. Sci. 3 (12) (2022) 101154, http://dx.doi.org/10.1016/j.xcrp.2022.101154.

[55] Z. Cui, N. Cui, J. Rao, C. Li, C. Zhang, Current distribution estimation of parallel-connected batteries for inconsistency diagnosis using long short-term memory networks, IEEE Trans. Transp. Electrif. 8 (1) (2022) 1013–1025, http://dx.doi.org/10.1109/TTE.2021.3118691.

[56] A. Reiter, S. Lehner, O. Bohlen, D.U. Sauer, Electrical cell-to-cell variations within large-scale battery systems — A novel characterization and modeling approach, J. Energy Storage 57 (2023) 106152, http://dx.doi.org/10.1016/j.est.2022.106152.

[57] A. Fill, T. Mader, T. Schmidt, A. Avdyli, M. Kopp, K.P. Birke, Experimental investigations on current and temperature imbalances among parallel-connected lithium-ion cells at different thermal conditions, J. Energy Storage 51 (2022) 104325, http://dx.doi.org/10.1016/j.est.2022.104325.

[58] R. Luca, M. Whiteley, T. Neville, T. Tranter, J. Weaving, J. Marco, P.R. Shearing, D.J.L. Brett, Current imbalance in parallel battery strings measured using a hall-effect sensor array, Energy Technol. 9 (4) (2021) 2001014, http://dx.doi.org/10.1002/ente.202001014.

[59] M. Al-Amin, A. Barai, T. Ashwin, J. Marco, An insight to the degradation behaviour of the parallel connected lithium-ion battery cells, Energies 14 (16) (2021) 4716, http://dx.doi.org/10.3390/en14164716.

[60] A. Fill, T. Schmidt, T. Mader, R. Llorente, A. Avdyli, B. Mulder, K.P. Birke, Influence of cell parameter differences and dynamic current stresses on the current distribution within parallel-connected lithium-ion cells, J. Energy Storage 32 (2020) 101929, http://dx.doi.org/10.1016/j.est.2020.101929.

[61] M. Ye, X. Song, R. Xiong, F. Sun, A novel dynamic performance analysis and evaluation model of series-parallel connected battery pack for electric vehicles, IEEE Access 7 (2019) 14256–14265, http://dx.doi.org/10.1109/ACCESS.2019.2892394.

[62] G. Piombo, J. Marco, R. Boyd, Analysis of parallel connected lithium-ion cells imbalanced performance based on electrothermal modelling environment, in: 2021 IEEE Vehicle Power and Propulsion Conference, VPPC, IEEE, 2021, pp. 1–7, http://dx.doi.org/10.1109/VPPC53923.2021.9699370.

[63] J. Antony, Design of Experiments for Engineers and Scientists, Elsevier, 2003, pp. 1–152, http://dx.doi.org/10.1016/B978-0-7506-4709-0.X5000-5.

[64] L.A. Román-Ramírez, J. Marco, Design of experiments applied to lithium-ion batteries: A literature review, Appl. Energy 320 (2022) 119305, http://dx.doi.org/10.1016/J.APENERGY.2022.119305.

[65] R. Rangappa, S. Rajoo, Effect of thermo-physical properties of cooling mass on hybrid cooling for lithium-ion battery pack using design of experiments, Int. J. Energy Environ. Eng. 10 (1) (2019) 67–83, http://dx.doi.org/10.1007/s40095-018-0284-6.

[66] J. E, D. Han, A. Qiu, H. Zhu, Y. Deng, J. Chen, X. Zhao, W. Zuo, H. Wang, J. Chen, Q. Peng, Orthogonal experimental design of liquid-cooling structure on the cooling effect of a liquid-cooled battery thermal management system, Appl. Therm. Eng. 132 (2018) 508–520, http://dx.doi.org/10.1016/j.applthermaleng.2017.12.115.

[67] D. Montgomery, Design and Analysis of Experiments, John Wiley and Sons Ltd, 2017, URL https://www.wiley.com/en-us/Design+and+Analysis+of+Experiments%2C+10th+Edition-p-9781119492443.

[68] B. Wu, W.D. Widanage, S. Yang, X. Liu, Battery digital twins: Perspectives on the fusion of models, data and artificial intelligence for smart battery management systems, Energy AI 1 (2020) 100016, http://dx.doi.org/10.1016/j.egyai.2020.100016.

[69] M. Aykol, P. Herring, A. Anapolsky, Machine learning for continuous innovation in battery technologies, Nat. Rev. Mater. 5 (10) (2020) 725–727, http://dx.doi.org/10.1038/s41578-020-0216-y.

[70] Z.C. Lipton, The mythos of model interpretability, Commun. ACM 61 (10) (2018) 36–43, http://dx.doi.org/10.1145/3233231.

[71] V. Belle, I. Papantonis, Principles and practice of explainable machine learning, Front. Big Data 4 (2021) 39, http://dx.doi.org/10.3389/fdata.2021.688969.

[72] C. Molnar, Interpretable Machine Learning, 2022, URL https://christophm.github.io/interpretable-ml-book.

[73] M. Faraji Niri, C. Reynolds, L.A. Román Ramírez, E. Kendrick, J. Marco, Systematic analysis of the impact of slurry coating on manufacture of li-ion battery electrodes via explainable machine learning, Energy Storage Mater. 51 (2022) 223–238, http://dx.doi.org/10.1016/j.ensm.2022.06.036.

[74] M. Faraji-Niri, K. Liu, G. Apachitei, L.A. Román-Ramírez, M. Lain, D. Widanage, J. Marco, Quantifying key factors for optimised manufacturing of li-ion battery anode and cathode via artificial intelligence, Energy AI 7 (2022) 100129, http://dx.doi.org/10.1016/j.egyai.2021.100129.

[75] M.J. Brand, P. Berg, E.I. Kolp, T. Bach, P. Schmidt, A. Jossen, Detachable electrical connection of battery cells by press contacts, J. Energy Storage 8 (2016) 69–77, http://dx.doi.org/10.1016/j.est.2016.09.011.

[76] H. Akaike, Information theory and an extension of the maximum likelihood principle, in: Biogeochemistry, Vol. 1998, Springer Science and Business Media Deutschland GmbH, 1998, pp. 199–213, http://dx.doi.org/10.1007/978-1-4612-1694-0_15.

[77] D.E. Coleman, D.C. Montgomery, A systematic approach to planning for a designed industrial experiment, Technometrics 35 (1) (1993) 1–12, http://dx.doi.org/10.1080/00401706.1993.10484984.

[78] M.J. Brand, E.I. Kolp, P. Berg, T. Bach, P. Schmidt, A. Jossen, Electrical resistances of soldered battery cell connections, J. Energy Storage 12 (2017) 45–54, http://dx.doi.org/10.1016/j.est.2017.03.019.

[79] L. Spitthoff, P.R. Shearing, O.S. Burheim, Temperature, ageing and thermal management of lithium-ion batteries, Energies 14 (5) (2021) 1248, http://dx.doi.org/10.3390/en14051248.

[80] F. Naseri, C. Barbu, T. Sarikurt, Optimal sizing of hybrid high-energy/high-power battery energy storage systems to improve battery cycle life and charging power in electric vehicle applications, J. Energy Storage 55 (2022) 105768, http://dx.doi.org/10.1016/J.EST.2022.105768.

[81] K.S. Ng, C.-S. Moo, Y.-P. Chen, Y.-C. Hsieh, Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries, Appl. Energy 86 (9) (2009) 1506–1511, http://dx.doi.org/10.1016/j.apenergy.2008.11.021.

[82] C.M. Anderson-Cook, C.M. Borror, D.C. Montgomery, Response surface design evaluation and comparison, 2009, http://dx.doi.org/10.1016/j.jspi.2008.04.004.

[83] StatEase, Design-expert, 2022, URL https://www.statease.com/software/design-expert/.

[84] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, F. Herrera, Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, Inf. Fusion 58 (2020) 82–115, http://dx.doi.org/10.1016/j.inffus.2019.12.012.

[85] S.M. Lundberg, P.G. Allen, S.-I.I. Lee, A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst. 30 (2017) 4766–4775, http://dx.doi.org/10.48550/arXiv.1705.07874.

[86] T. Hastie, R. Tibshirani, J. Friedman, the Elements of Statistical Learning, in: Springer Series in Statistics, Springer New York, New York, NY, 2009, http://dx.doi.org/10.1007/978-0-387-84858-7.

[87] A. Goldstein, A. Kapelner, J. Bleich, E. Pitkin, Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation, J. Comput. Graph. Statist. 24 (1) (2015) 44–65, http://dx.doi.org/10.1080/10618600.2014.907095.

[88] L.S. Shapley, A Value for N-Person Games, Princeton University Press Princeton, 1953, http://dx.doi.org/10.1515/9781400881970-018.

[89] A.B. Owen, C. Prieur, On Shapley value for measuring importance of dependent inputs, SIAM/ASA J. Uncertain. Quantif. 5 (1) (2017) 986–1002, http://dx.doi.org/10.1137/16M1097717.

[90] E. Song, B.L. Nelson, J. Staum, Shapley effects for global sensitivity analysis: Theory and computation, SIAM/ASA J. Uncertain. Quantif. 4 (1) (2016) 1060–1083, http://dx.doi.org/10.1137/15M1048070.

[91] G. Pozzato, A. Allam, S. Onori, Lithium-ion battery aging dataset based on electric vehicle real-driving profiles, Data Brief 41 (2022) 107995, http://dx.doi.org/10.1016/j.dib.2022.107995.

[92] E. Catenaro, S. Onori, Experimental data of lithium-ion batteries under galvanostatic discharge tests at different rates and temperatures of operation, Data Brief 35 (2021) 106894, http://dx.doi.org/10.1016/j.dib.2021.106894.

[93] S. Ha, G. Pozzato, S. Onori, Electrochemical characterization tools for lithium-ion batteries, J. Solid State Electrochem. (2023) 1–27, http://dx.doi.org/10.1007/s10008-023-05717-1.

[94] LGChem, Product specification: Rechargeable lithium ion battery. Model: INR21700 M50t 18.2Wh, 2018, https://www.batteryspace.com/prod-specs/11514.pdf.

[95] Samsung, Specification of product: Lithium-ion rechargeable cell. Model: INR21700-50e, 2019, https://www.batteryspace.com/prod-specs/7015-INR21700-50E.pdf.

[96] G. Piombo, S. Fasolato, R. Heymer, M. Hidalgo, M. Faraji Niri, M.D. Raimondo, J. Marco, S. Onori, Under review: Full factorial design of experiments dataset for parallel-connected lithium-ion cells imbalanced performance investigation, Data Brief (2024).

[97] M.O. Akinwande, H.G. Dikko, A. Samson, Variance inflation factor: As a condition for the inclusion of suppressor variable(s) in regression analysis, Open J. Stat. 05 (07) (2015) 754–767, http://dx.doi.org/10.4236/OJS.2015.57075.

[98] T.F. Cootes, M.C. Ionita, C. Lindner, P. Sauer, Robust and accurate shape model fitting using random forest regression voting, in: Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VII 12, Springer, 2012, pp. 278–291, http://dx.doi.org/10.1007/978-3-642-33786-4_21.