

# Efficient Analysis of Interdependencies in Electrode Manufacturing Through Joint Application of Design of Experiments and Explainable Machine Learning

SajedeH Haghi,<sup>\*[a]</sup> Josef Keilhofer,<sup>[a]</sup> Nico Schwarz,<sup>[a]</sup> Pengdan He,<sup>[a]</sup> and Rüdiger Daub<sup>[a]</sup>

Battery cell production is a key contributor to achieving a net-zero future. A comprehensive understanding of the various process steps and their interdependencies is essential for speeding up the commercialization of newly developed materials and optimizing production processes. While several approaches have been employed to analyze and understand the complexity of the process chain and its interdependencies – ranging from expert- and simulation-based to data-driven – the latter holds significant potential for real-time application. This is particularly relevant for inline process control and optimization. To streamline the development and implementation of data-driven models, a holistic framework that encompasses all necessary steps – from identification of relevant parameters and

generation of data to development of models – is imperative. This article aims to address this objective by presenting a comprehensive and systematic methodology, demonstrated for efficient cross-process analysis in electrode manufacturing. Through the combined utilization of design of experiments methods, data-driven models, and explainable machine learning methods, the interdependencies between production parameters and the physical, mechanical, and electrochemical characteristics of the electrodes were uncovered. These actionable insights are essential for enabling informed decision-making, facilitating the selection of appropriate process parameters, and ultimately optimizing the production process.

## Introduction


Over the last decade, substantial progress has been made within the lithium-ion battery field. These advancements have been focused on both materials and production, with the overarching goal of improving quality and reducing cost.<sup>[1]</sup> However, there are still certain challenges that need to be addressed to ensure the widespread adoption of electric vehicles in the market and achieve the vision of a sustainable society.


From the production perspective, one of the key challenges lies in the complexity of the process chain, resulting from multiple interrelated variables.<sup>[2]</sup> Given the intricate nature of the process chain and its numerous interdependencies, the causality between manufacturing parameters and the resulting performance of both the electrode and the final cell remains largely unknown. A comprehensive understanding of process steps, the parameters, their interactions, and aggregated effects on the intermediate and final product quality provides the foundation for event-driven decisions and pursuing holistic

optimization strategies.<sup>[3]</sup> Within battery cell production, electrode manufacturing is of particular importance, as the majority of cell properties are established in this phase.<sup>[4]</sup>

With the rise of digitalization and the contemporary industrial revolution, data-driven approaches have become the driving force for a paradigm shift in tackling the complexity of the process chain and gaining an in-depth process understanding.<sup>[5,6]</sup> While several studies have showcased the potential of data-driven approaches in battery cell production over the last five years,<sup>[7–11]</sup> there are still certain aspects that need to be addressed to accelerate the development and adoption of these methods, ultimately paving the way toward the realization of a smart battery cell manufacturing vision. In a previous publication, a comprehensive mapping study was conducted, outlining the process steps and interdependencies analyzed using data-driven models in battery cell production.<sup>[6]</sup> The study highlights both the domain-specific and overarching aspects that have received limited attention from the battery production research community. The results indicate that certain process steps, such as drying in electrode manufacturing, have not been sufficiently investigated, specifically at the pilot scale level. Moreover, most studies focused on individual process steps, with only a limited number of articles exploring cross-process effects.<sup>[6]</sup> In the case of the latter, the effects have not been restricted to immediate consecutive process steps; for instance, some studies analyzed the interdependencies between the coating and calendaring processes, under the premise that the effect of the intervening process, in this case, the drying step, can be effectively isolated or disregarded. Overlooking the impacts of an intervening process when evaluating cross-process effects can induce a certain degree of

[a] S. Haghi, J. Keilhofer, N. Schwarz, P. He, Prof. Dr. R. Daub  
Institute for Machine Tools and Industrial Management  
Technical University of Munich  
Boltzmannstr. 15, Garching 85748 (Germany)  
E-mail: sajedeH.haghi@iwb.tum.de

 Supporting information for this article is available on the WWW under <https://doi.org/10.1002/batt.202300457>

 © 2023 The Authors. Batteries & Supercaps published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

uncertainty in the developed models. This is especially the case in electrode manufacturing, where significant interdependencies of subsequent processes are apparent. From the material system perspective, the majority of the studies have focused on cathodes.<sup>[6]</sup>

LOMBARDO ET AL.<sup>[5]</sup> reviewed the application of machine learning (ML) methods in battery research, covering material characterization, manufacturing, and diagnosis. The authors have highlighted the importance of the first step toward the development of data-driven models, specifically the selection of parameters to be analyzed, suitable acquisition methods, data quality, and veracity.<sup>[5]</sup> In a preceding publication, HAGHI ET AL.<sup>[12]</sup> introduced a two-step approach, based on a literature review and expert insights, to evaluate parameters concerning their relevance in electrode manufacturing. The results were based on the assessment of two dimensions: the importance of the parameter, evaluated using the number of interdependencies and their impact on the intermediate or final product quality, and the degree of complexity associated with the digitalization of the parameter. In a subsequent work,<sup>[13]</sup> an overview of possible acquisition methods for the characterization of intermediate products in electrode manufacturing was presented. The intermediate products serve as quality indicators for both the set process parameters and the final product quality. Building upon the insights provided in the previous publications,<sup>[6,12,13]</sup> this article aims to proceed with the steps toward data generation and the development of data-driven models.

Design of Experiments (DoE) can be considered as a complementary methodology for a comprehensive process analysis using data-driven models.<sup>[14]</sup> The DoE methods facilitate efficient analysis by generating an informative, statistically reliable dataset while minimizing the experimental efforts.<sup>[5,14]</sup> In a comprehensive review, ROMÁN-RAMÍREZ & MARCO<sup>[15]</sup> provided an overview of DoE methods, their objectives, and their application in lithium-ion battery research. For optimization studies and predictive models, the response surface methodology (RSM) is suggested.<sup>[15,16]</sup> Among various RSM methods, the I-optimal design stands out, particularly for prediction purposes, as it provides a comprehensive understanding of the

input variables' impacts on the responses, both quantitatively and qualitatively. Additionally, the method can pinpoint the optimum regions and settings, while minimizing the average prediction variance across the experimental space.<sup>[15,17,18]</sup> Despite these distinct benefits, this method has not been widely adopted in battery cell production research, particularly in combination with data-driven models.

This article seeks to address the existing research gaps by demonstrating the application of RSM methods for efficient data generation, development of data-driven models and revealing the interdependencies between various process steps. The remainder of this article is structured as follows. The next section provides an overview of the systematic approach adopted for DoE. This is followed by a detailed description of the experimental procedures involved in the production and characterization of the electrodes. The focus then shifts to data preprocessing and analysis, while the results derived from the developed data-driven models are presented in the consequent section. Finally, concluding remarks and an outlook on further research activities are provided.

## Systematic Design of Experiments

Considering the existing literature investigating the formulation aspect,<sup>[6]</sup> this study narrows its focus to a single formulation and slurry, analyzing the effects of the coating, drying, and calendaring processes. The results from the previous publication<sup>[12]</sup> were adopted to select the quality-relevant parameters to be considered in the DoE. Given the technology available at the pilot production line of the Institute for Machine Tools and Industrial Management at the Technical University of Munich – specifically, a roll-to-roll doctor blade coating machine – the list of parameters was adjusted. Table 1 provides an overview of all intermediate product parameters evaluated as highly important, in both the literature-based approach and the expert interviews,<sup>[12]</sup> and their complexity regarding digitalization.

While the topic of defects in both the coating process – primarily caused by slot-die coating or inhomogeneous slurry –

**Table 1.** Overview of quality-relevant product parameters and their complexity concerning digitalization based on the findings of Ref. [12].

Process	Product parameter	Importance	Complexity
Coating	Wet film thickness	High	Low
	Quality of wet film	High	High
Drying	Adhesion	High	High
	Electrode thickness	High	Low
	Mass loading	High	Low
	Porosity	High	High
	Quality of electrode (defects)	High	High
Calendaring	Electrode thickness	High	Low
	Porosity	High	High
	Adhesion	High	High
	Quality of electrode (defects)	High	High

**Table 2.** Overview of considered parameters as factors and their range for the I-optimal DoE.

Process	Factor	Range
Coating	Mass loading* [mg/cm <sup>2</sup> ]	8.2–11.2
Drying	Temperature of the second dryer [°C]	50–65
Calendering	Porosity [%]	25–40

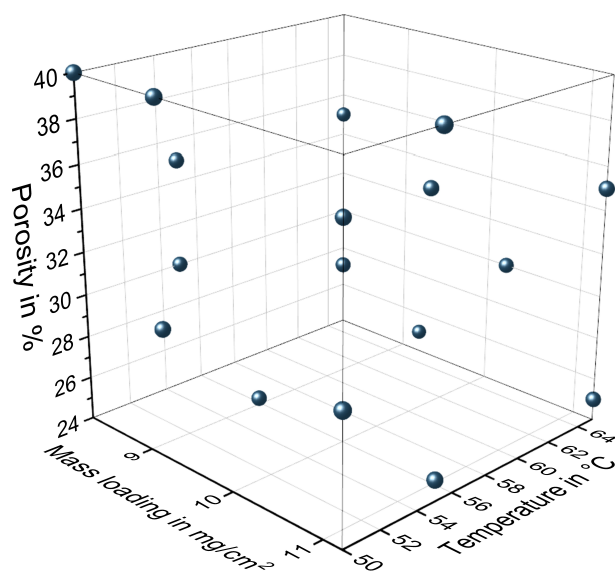
\* Range reported for Active Material (AM)

and the calendering process is noteworthy, these parameters were not considered for the analysis in the DoE. This decision was driven by various factors, including the need to limit the number of experiments to underscore the benefits of DoE for cross-process analysis, the available coating technology using a doctor blade, along comprehensive reviews available in the literature concerning defects.<sup>[19,20]</sup> To ensure an exhaustive cross-process analysis with independent input variables, the electrode porosity after calendering and the mass loading were chosen as input variables, eliminating the third dependent variable, the thickness. Due to the complexity involved in the digitalization of adhesion, with only offline characterization methods available,<sup>[13]</sup> and the difficulty in defining appropriate levels for the DoE, the temperature of the second dryer was chosen as an influential factor impacting adhesion.<sup>[21,22]</sup> Given the linear correlation between the wet film thickness and the mass loading, only the latter was considered a relevant product parameter representing the coating and drying process. Table 2 summarizes the considered factors and their respective ranges for the I-optimal DoE. The selected ranges for mass loading and porosity align with those commonly explored in the literature, while the range for the drying temperature was chosen based on the domain know-how.

The experimental plan was generated with the help of the Design-Expert® software. Figure 1 shows the experimental space consisting of 17 different configurations. The Variance Inflation Factor (VIF) was used to confirm the independency of the analyzed parameters and ensure that no substantial multicollinearity exists.<sup>[23]</sup> It should be noted that for a conventional full factorial design, based on the specified factors and levels, a total of 64 experimental runs would be needed. By leveraging the DoE optimized methods, this requirement can be drastically reduced to 17 runs. This approach not only streamlines a holistic analysis but also significantly reduces the cost and experimental efforts.

## Experimental Section

Following the generated experimental plan, 17 electrode configurations were produced and characterized. Though the primary focus of this study was on the anode manufacturing, it was essential to replicate genuine anode behavior at the cell level. To ensure this, corresponding cathodes were manufactured with an N:P ratio of 1.2 and a porosity of 30%. The cathodes were subsequently used to assemble full CR2032 coin cells, conduct the electrochemical characterization and evaluate the overall cell



**Figure 1.** The experimental space designed using the I-optimal method, consisting of 17 different electrode configurations with variations in porosity, active mass loading and drying temperature.

performance. All the models developed were based on the calendered electrodes. However, for a comprehensive understanding, the uncalendered electrodes, consisting of 13 distinct mass loadings and drying temperatures, were additionally characterized using the adhesion strength test and electrochemical impedance spectroscopy. The results are presented in the data preprocessing and analysis section.

## Electrode Manufacturing

### Slurry Mixing

To prepare the anode slurry, a dry mixture consisting of graphite (SMGA5, Showa Denko, Japan) and carbon black (Super C-65, Imerys, Switzerland) was first blended at 1400 rpm for 1 min using a centrifugal mixer (Speedmixer DAC 1100.2 VAC-P, Hauschild, Germany). Separately, distilled water and sodium carboxymethyl cellulose (CMC,  $M_w \approx 250,000$  g mol, Merck, Germany) were pre-mixed at 2000 rpm for 40 min using a disperser (FM, 10-SIP, VMA-Getzmann, Germany). Subsequently, the dry mixture was added to the water and CMC blend in five equal portions, with dispersion continued for an additional 45 min at 2000 rpm. Lastly, styrene-butadiene rubber (SBR, Zeon, Japan) was added and mixed in for 20 minutes at 500 rpm.

For cathode slurry, lithium-nickel-cobalt-manganese-oxide (HED™ NCM-622 DT011, BASF, Germany), carbon black (Super C-65, Imerys, Switzerland), and conductive graphite (C-Nergy SFG6L, Imerys, Switzerland) were initially blended using the centrifugal mixer at 1400 rpm for 1 min. Concurrently, N-Methyl-2-pyrrolidone (NMP 328634, Merck, Germany) and polyvinylidene difluoride (PVDF Solef 5130, Solvay, Belgium) were blended using the disperser for 40 min at 2000 rpm. Following this, the pre-blended NCM with carbon black and conductive graphite was added in five equal portions to the NMP and PVDF mixture and dispersed for an additional 45 min at 2000 rpm. Both slurries were then degassed in the centrifugal mixer at an absolute pressure of 250 mbar, running at 600 rpm for 5 min. The rheological characteristics of the produced slurries were determined using a rheometer (Kinexus pro, Malvern Panalytical, Germany) with a cone-plate configuration, measuring the viscosity

versus the shear rate. At 25 °C and a shear rate of 100 s<sup>-1</sup>, the anode slurry exhibited a viscosity of 2.6 Pa·s, while the viscosity of the cathode slurry was 3.6 Pa·s. Table 3 summarizes the formulations and relevant information for both anode and cathode slurries.

### Coating and Drying

The anode slurry was continuously coated on a copper current collector foil with a thickness of 10 μm (SE-Cu58, Schlenk, Germany) using a roll-to-roll coating machine equipped with a doctor blade coating technology and three infrared dryers (BC50, Coatema, Germany). The temperature of the second dryer was varied, as described in the DoE section. The temperatures for the first and third dryers were consistently set at 50 °C and 65 °C, respectively with a coating speed of 0.6 m/min. The wet film mass loading, immediately after coating, and the dry film mass loading were measured inline using ultrasound systems (CFS400-USMX200 and OF400-USMX200, MeSys, Germany). The electrodes were produced based on the target mass loading with a tolerance of 5%. In addition to the inline mass loading measurements, the thickness of the produced electrodes was measured based on samples using a digital micrometer (40 EWRi, Mahr, Germany).

The cathode slurry was coated on an aluminum current collector foil with a thickness of 15 μm (AA1100/H19, Speira, Germany). The drying temperatures of the three dryers were set to 60 °C, 80 °C and 100 °C, respectively. In correspondence with the defined mass loading based on the active material for the anodes, as detailed in the DoE section, four different cathodes with active mass loadings of 14.27, 16.01, 17.75, and 19.49 mg/cm<sup>2</sup> were produced. The first two configurations, with relatively lower mass loadings, were dried at a faster speed of 0.7 m/min speed, while the last two configurations were dried at 0.5 m/min. The rationale for choosing different drying speeds for these cathode configurations stems from the aim of fine-tuning the drying process and ensuring a consistent electrode quality across varying mass loadings.

### Calendering

The produced electrodes were subsequently trimmed to A4-sized sheets before calendering and characterized based on the mass and thickness. The mass was measured using an analytical balance (AX26 Comparator, Mettler Toledo, Switzerland). The porosity  $\varepsilon$  was calculated based on the bulk density of the electrode and its component following the equation below:

$$\varepsilon = 1 - \frac{\rho_{\text{bulk}}}{\rho_{\text{theor.}}} = 1 - \frac{m_{\text{coating}}}{A_{\text{coating}} \cdot d_{\text{coating}} \cdot \rho_{\text{theor.}}} \quad (1)$$

**Table 3.** Summary of formulations and relevant information for the produced electrodes.

	Anode	Cathode
Active material	Graphite 94 wt%	NCM622 95.5 wt%
Conductive additive	C65 1 wt%	C65 & SFG6L 2.25 wt% & 0.75 wt%
Binder additive	CMC & SBR 2 wt% & 3 wt%	PVDF 1.5 wt%
Theoretical density	2.158 g/cm <sup>3</sup>	4.447 g/cm <sup>3</sup>
Solid content	52 wt%	78 wt%
Specific capacity	355 mAh/g	170 mAh/g

In Equation (1), the term  $\rho_{\text{bulk}}$  denotes the electrode's bulk density, which is derived from its mass  $m_{\text{coating}}$  and volume, calculated using the electrode's area  $A_{\text{coating}}$  and thickness  $d_{\text{coating}}$ .  $\rho_{\text{theor.}}$  represents the theoretical density of the electrode's components (see Table 3). Based on Equation (1) and the target porosities defined in the experimental plan, the target electrode thicknesses for different configurations were calculated. In the case of the cathodes, the target porosity was set to 30%. To achieve the specified thicknesses, the electrodes were processed using a calender with a roll diameter of 400 mm (EA102, Coatema, Germany). The calendering gap was primarily used to achieve the target electrode thickness, while the roll pressure was adjusted as needed for fine-tuning the process. A constant web speed of 1 m/min was chosen for the calendering. The anodes were calendered at room temperature, whereas for the cathodes the rolls were heated to 100 °C.

### Coin Cell Assembly

Symmetric and full coin cells were used for electrochemical characterization of the produced electrodes. The electrodes, separator, and cell components were dried at 120 °C for at least 12 hours in a vacuum dryer (Goldbrunn 1450, Goldbrunn, Germany) to remove residual moisture. The coin cells were assembled in a dry room with a minimum dew point of -40 °C. For the full coin cells, a 15 mm diameter anode was separated from a 14 mm diameter cathode using a 16 mm diameter glass fiber separator (Type 691, VWR, USA). The electrodes were assembled with metal spacers, with a total thickness of 1.1 mm. For the symmetric cells, the total thickness of the spacers was 1.5 mm. For the full coin cells, 100 μL electrolyte (LP572, BASF, Germany) was used. This electrolyte is a blend of ethylene carbonate (EC) and ethyl methyl carbonate (EMC) at a mixing ratio of 3:7, supplemented with 1 M lithium hexafluorophosphate (LiPF<sub>6</sub>) conductive salt and 2 wt% vinylene carbonate (VC). For symmetric cells, 100 μL of a non-intercalating electrolyte was used to ensure blocking conditions. This electrolyte comprises 0.01 M tetrabutylammonium perchlorate dissolved in an EC:EMC mixture at a volume fraction of 3:7.<sup>[24]</sup> An electrolyte conductivity of 350 μS/cm was measured using a laboratory conductivity sensor (InLab, Mettler Toledo, Switzerland). To ensure that the measurement is representative of the symmetric cell condition, both the sensor and the non-intercalating electrolyte were pre-conditioned in a heating chamber at 25 °C for 3 hours prior to the measurement.

### Mechanical Characterization

The electrode's mass loading and thickness were measured as metrics representing the electrode's physical characteristics. Additionally, the produced anodes were characterized mechanically through adhesion strength using a pull-off test performed on a material testing machine (Z050, ZwickRoell, Germany). The test was largely based on the procedure presented by HASELRIEDER ET AL.,<sup>[25]</sup> with certain adjustment described in the following. A double-sided adhesive tape (5696 extra strong, Tesa SE, Germany) was used for the procedure. To achieve a uniform areal force distribution, a round specimen was chosen over a rectangular one. The size of the specimen and the compression stress were adjusted to ensure a high degree of removal (see Table 4).

### Electrochemical Characterization

The symmetric cells were used to investigate the ionic resistance and tortuosity of the anodes with the help of electrochemical

Table 4. Relevant parameters for adhesion strength measurements.	
Parameter	Value
Data acquisition rate [Hz]	2000
Compression stress [kPa]	1500
Specimen size [mm <sup>2</sup> ]	50.26
Dwell time [s]	30
Compression velocity [mm/min]	0.75
Pull-off velocity [mm/min]	100

impedance spectroscopy (EIS).<sup>[26,27]</sup> The full coin cells went through formation and discharge rate capability test.

### Potentiostatic Electrochemical Impedance Spectroscopy

The EIS measurements were conducted at a temperature of 25 °C using a potentiostat (VSP-3e, BioLogic, France). Across a frequency range of 10 kHz to 100 mHz, ten logarithmically spaced measurement points per decade were recorded. During the tests, an alternating excitation voltage with an amplitude of 20 mV around the open circuit voltage (OCV) was applied.

### Formation and Discharge Rate Capability Test

The formation and cell rate capability test were conducted at 25 °C in a battery test system (CTS, BaSyTec, Germany). Cell capacities were determined based on the last discharge cycle during the formation process. The formation began with two charge and discharge cycles at a constant current (CC) of 0.2, ranging between 2.9 V and 4.2 V. The third cycle used a CC charge at 0.2 C, transitioning to a constant voltage (CV) charge at 4.2 V until the current fell below 0.02 C. The final discharge cycle was carried out at a CC of 0.2 C. For the discharge rate capability test, cells were discharged at varying CC-rates, from 0.1 C up to 5 C, until the voltage decreased to 2.9 V. The detailed protocols are available in the supplementary section.

### Data Preprocessing and Analysis

During the coating and drying process, process and intermediate product parameters, such as coating gap, wet and dry mass loading, were collected inline. Additionally, the electrode thickness was measured offline and documented based on samples. Given the process's nature, with varied sensors located at specific positions over a span of approximately 6 meters, data points had to be systematically mapped. For this purpose, a script was developed to automatically map the coating gap to its corresponding wet mass loading, dry mass loading, and thickness, considering the distance between the sensors and the coating speed. It should be noted that all 13 different configurations were coated on a single coil. The results additionally included segments of the coated coil that were dried at specific temperatures according to the experimental plan.

Multiple samples for each configuration, including the non-calendered electrodes, were used for the adhesion measurement. The recorded data was cleaned, and outliers, i.e. values falling outside of  $\pm 1.5$  standard deviations from the mean

value, were excluded. The final dataset included 51 measurements for the calendered electrodes and 39 for the uncalendered ones. Figure 2 presents a contour plot, generated using *Origin* software, visualizing the adhesion response in relation to both mass loading and drying temperature for the uncalendered electrodes. At the lowest mass loading of 8.2 mg/cm<sup>2</sup>, a low temperature correlates with high adhesion strength. For this mass loading, changes in temperature above 55 °C did not significantly affect the adhesion strength. This might be attributed to the pore emptying phase<sup>[22]</sup> occurring in the first dryer for thinner electrodes in the conducted study. For intermediate mass loadings, the data suggests a local optimum for drying temperature, falling between 55 to 60 °C. Notably, the combination of the highest mass loading and the highest temperature leads to poor adhesion, which is consistent with the findings from the literature.<sup>[21]</sup>

The data collected from the EIS measurements were analyzed and fitted to a Transmission Line Model (TLM) under blocking conditions using the *impedance.py* package,<sup>[28]</sup> as reported by LANDESFEIND ET AL.<sup>[27,29]</sup> In the first step, the Nyquist plots of the impedance spectra were visually inspected to identify measurements that did not align with the TLM model. Subsequently, outliers were identified based on the values of ionic resistance, using the same approach described for the adhesion data. The symmetric cells were produced for both calendered and uncalendered electrodes. Following data cleaning, 86 data points were retained for the calendered electrodes and 51 data points for the 13 different configurations of uncalendered electrodes, resulting in a total of 137 data points. Based on the literature, it was expected that calendaring has a significant influence on the contact resistance. A pre-analysis of the Nyquist plots revealed an increased contact resistance appearing in the high-frequency range for uncalendered electrodes.<sup>[30]</sup> Hence, the frequency range for the analysis of the uncalendered electrodes was limited to 1 kHz–100 mHz. Within this range, the imaginary contribution of the contact resistance vanishes, allowing a more accurate estimation of the ionic resistance. Figure 3 provides an overview of the ionic resistance

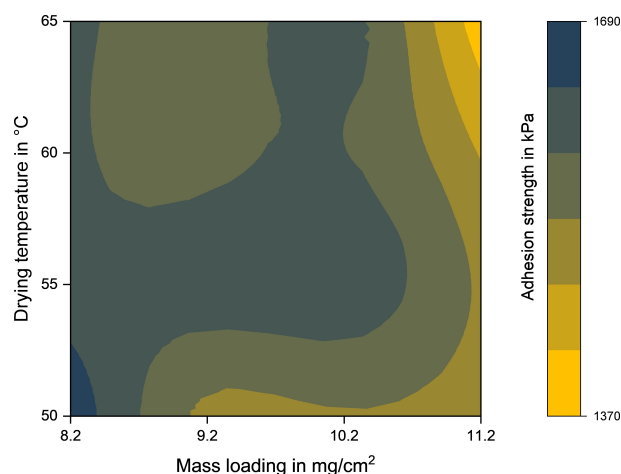
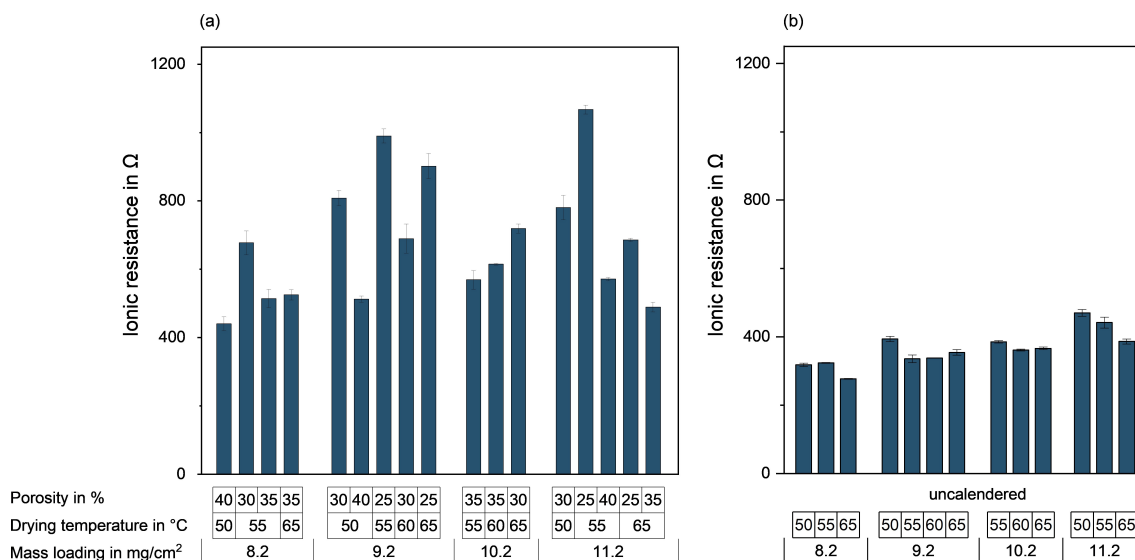


Figure 2. Contour plot: Adhesion response to mass loading and drying temperature for uncalendered electrodes.



**Figure 3.** Ionic resistance based on impedance spectroscopy data: a) calendered electrodes, b) uncalendered electrodes.

of the produced anodes. The data indicates that a higher mass loading leads to an increased ionic resistance. Moreover, a low porosity results in greater ionic resistance. When considering configurations with identical mass loadings and porosities but varying drying temperatures, ionic resistance decreases as the temperature rises. This effect is particularly noticeable for anodes with higher mass loading, evident in both calendered and uncalendered electrodes. While the effects of mass loading and porosity in combination with the ionic resistance are relatively well-known in the literature, the impact of drying temperature has received limited attention. MORASCH ET AL.<sup>[31]</sup> explored binder gradients resulting from extreme drying temperatures, ranging from room temperature to 125 °C, on a laboratory scale for NMP-based graphite anodes. Utilizing EIS, their study demonstrated that electrodes with significant binder migration exhibited an increased ionic resistivity toward the separator side. This observation is in contrast with the findings from the experimental investigations outlined in this study. Based on the adhesion analysis, as presented in the data preprocessing and analysis section, it should be noted that extreme binder migrations were not identified. However, elevated temperatures in combination with mass loading can impact the microstructure of the electrodes,<sup>[32]</sup> potentially leading to variations in pore structure. As a result, reduced ionic resistance was observed for electrodes dried at higher temperatures, a trend consistent across both calendered and uncalendered variants. Based on the ionic resistance data, the tortuosity  $\tau$  can be calculated using the following equation, as proposed by LANDESFEIND ET AL.<sup>[27]</sup>

$$\tau = \frac{R_{\text{ion}} A \kappa \varepsilon}{2d} \quad (2)$$

In this equation, the ionic resistance  $R_{\text{ion}}$  is halved to account for the impedance introduced by the two identical electrodes in the symmetric cells. The term  $A$  represents the

cross-sectional area of the electrodes, while  $d$  and  $\varepsilon$  denote the average values for the thickness and porosity of the identical electrodes, respectively. The term  $\kappa$  represents the electrolyte conductivity.

For the electrochemical characterization conducted using full cells, a total of 51 data points was considered. The discharge capacities from the second cycle at various C-rates were selected for the model development. Additionally, the gravimetric capacity was determined based on the active mass of the electrode.

## Prediction of Product Properties Using Data-Driven Models

In this section, the results of the developed data-driven models based on the target variables are presented. Given the size of the dataset, Random Forest (RF) and Support Vector Machine (SVM), as the most commonly used algorithms in battery cell production,<sup>[6]</sup> were chosen. Considering the potential overfitting issues associated with RF, the eXtreme Gradient Boosting (XGBoost) modeling technique was adopted in certain cases. In addition to these advanced algorithms, the potentials of multiple linear regression (MLR) and polynomial regression were also investigated. A detailed explanation of the working principles of various ML algorithms and methods is beyond the scope of this article. For comprehensive descriptions, numerous publications and handbooks such as Refs. [5, 33–36] are suggested.

To provide insights into the complex models and uncover the multiple interdependencies, eXplainable Machine Learning (XML) techniques<sup>[37]</sup> were adopted. XML methods provide several compelling advantages. Primarily, they can foster increased trust in complex models, often referred to as *black box* models due to their opaque nature. By deploying XML methods, these black box models are converted into *glass box*

models, clarifying the rationale behind their decision-making processes. Moreover, the XML methods offer actionable insights, providing guidelines on how adjustments can be made based on the significance of specific parameters.<sup>[38,39]</sup> In a recent publication, FARAJI-NIRI ET AL.<sup>[40]</sup> reviewed the application of XML in the battery value chain. While the topic of ML has received increased attention over the last five years, the novel XML methods remain relatively unexplored, particularly in battery production.

In this article, the Feature Importance based on the Mean Decrease Impurity (MDI) as one of the common XML methods for tree-based models was adopted.<sup>[41–43]</sup> This method calculates the significance of each feature by evaluating the total reduction in the model's uncertainty or node impurity. This value is then weighted by the frequency with which data samples arrive at the respective decision point.<sup>[44]</sup> Additionally, the SHapley Additive exPlanations (SHAP) values were used to provide instance-level explanations.<sup>[45]</sup> Originating from cooperative game theory, SHAP is a method used to explain ML models by assigning an importance value, known as SHAP value, to each feature in the model. This method is known as a powerful tool for interpreting complex models, ensuring accuracy, robustness, and consistency in the explanations.<sup>[46]</sup> While the Feature Importance method using the MDI quantifies the overall relevance of the features, SHAP values not only indicate the relevance but also the direction in which a feature impacts the target variable.

For model development, all input features were normalized to ensure balanced variance and efficient training. In case of the SVM, given the working principle of the algorithm, the target variable was also normalized.<sup>[47]</sup> This is particularly relevant when the target variable encompasses a broad range, for example in case of the ionic resistance. The SVM aims to identify a hyperplane passing through as many data points as possible within a certain minimized distance. Without normalization, variations in data scales can distort this distance and potentially impact the model's accuracy. Further details regarding the selected hyperparameters for the developed models can be found in the supplementary information. The dataset was divided into an 80–20 split for training and testing, respectively, with the test data points randomly selected. When comparing different models aimed at predicting a specific target feature, the same test dataset was consistently employed to ensure the comparability. All data preprocessing, model development, and evaluation tasks were conducted using *Python 3.8*, a powerful programming language widely adopted in the fields of ML and data science. Specifically, the following libraries and packages were utilized: *pandas*, *numpy*, *matplotlib*, *seaborn*, *Scikit-learn*, *xgboost*, and *shap*.

For the evaluation of the developed models, three metrics were adopted. The r-squared ( $R^2$ ), also known as the coefficient of determination, quantifies how well the model explains the variance in the input variables.<sup>[48]</sup> With a range between 0 and 1, a value closer to 1 indicates a more accurate representation of the data by the model. To ensure the generalization of the model, the  $R^2$  was reported for both training and test dataset. The Root Mean Squared Error (RMSE) is another widely used

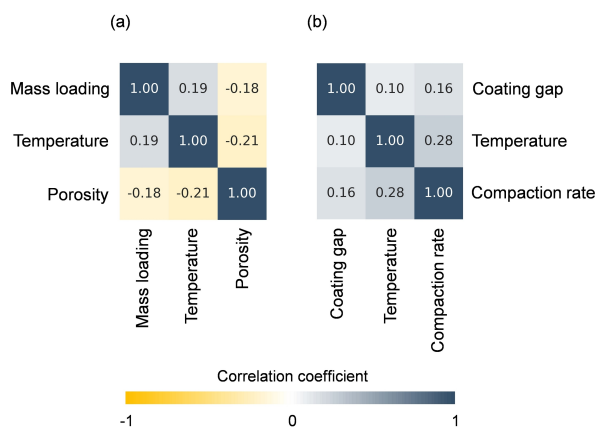
performance metric in statistics.<sup>[48]</sup> It measures the average magnitude of the errors between predicted and actual values, providing insight into the model's overall accuracy. A lower RMSE indicates a closer fit of the predicted values to the actual ones, while a higher RMSE suggests larger discrepancies between the predictions and true values. Additionally, it is important to examine the relationship between the Mean Squared Error (MSE) for both the training and test datasets. Ideally, these values should be of a similar magnitude.<sup>[36]</sup> A significantly larger MSE for the test dataset compared to the training dataset suggests overfitting, implying that while the model performs well on the training data, it may not generalize effectively to new, unseen data. The squared term places emphasis on larger errors, amplifying their impact.

In addition to the evaluation metrics, the parity plot was utilized as a visualization tool to showcase the performance of the developed models. The x-axis depicts the actual values or ground truth, and the y-axis displays the model's predicted values. In an ideal scenario with a perfect prediction, all data points would align with the 45° diagonal line.

The model development primarily revolved around three scenarios. In the first scenario, the input variables consisted of the sample-specific product parameters from the DoE and the drying temperature. The second one employed the corresponding process parameters, specifically coating gap, compaction rate, and drying temperature, as input variables. For more complex target values, such as adhesion, a third scenario was defined, incorporating both process parameters and intermediate product parameters collected during the processes. In the field of ML, particularly XML, it is crucial to ensure that input variables are independent. This requirement allows the model to differentiate between the distinct effects and evaluate the significance of each feature and its impact on the target variable. While the first two scenarios fulfill this aspect, due to the applied DoE, the third one does not. It is evident that in the last scenario the intermediate product parameters, such as wet mass loading, correlate with the corresponding process parameters, in this case, the coating gap. Hence, for this scenario, the XML methods were not adopted. It should be noted that the compaction rate was selected as a substitute for the calendaring gap as a process parameter. The choice was made due to the fact that the calendaring gap directly correlates with the mass loading and, subsequently, the coating gap. Based on the electrode thickness after drying  $d_{\text{uncal.}}$  and after calendaring  $d_{\text{cal.}}$ , the compaction rate  $\Pi$  is calculated as follows:<sup>[49]</sup>

$$\Pi = 1 - \frac{d_{\text{cal.}}}{d_{\text{uncal.}}} \quad (3)$$

Figure 4 showcases the correlation matrices, which include the coefficients for the input variables considered in the first two scenarios. Given that all correlation coefficients are below 0.3, the input variables exhibit weak correlations,<sup>[50]</sup> suggesting that they are suitable for the model development, application of XML methods and analysis of the interdependencies.



**Figure 4.** Exemplary correlation matrices for the considered input variables a) sample-specific product parameters and the drying temperature b) corresponding process parameters.

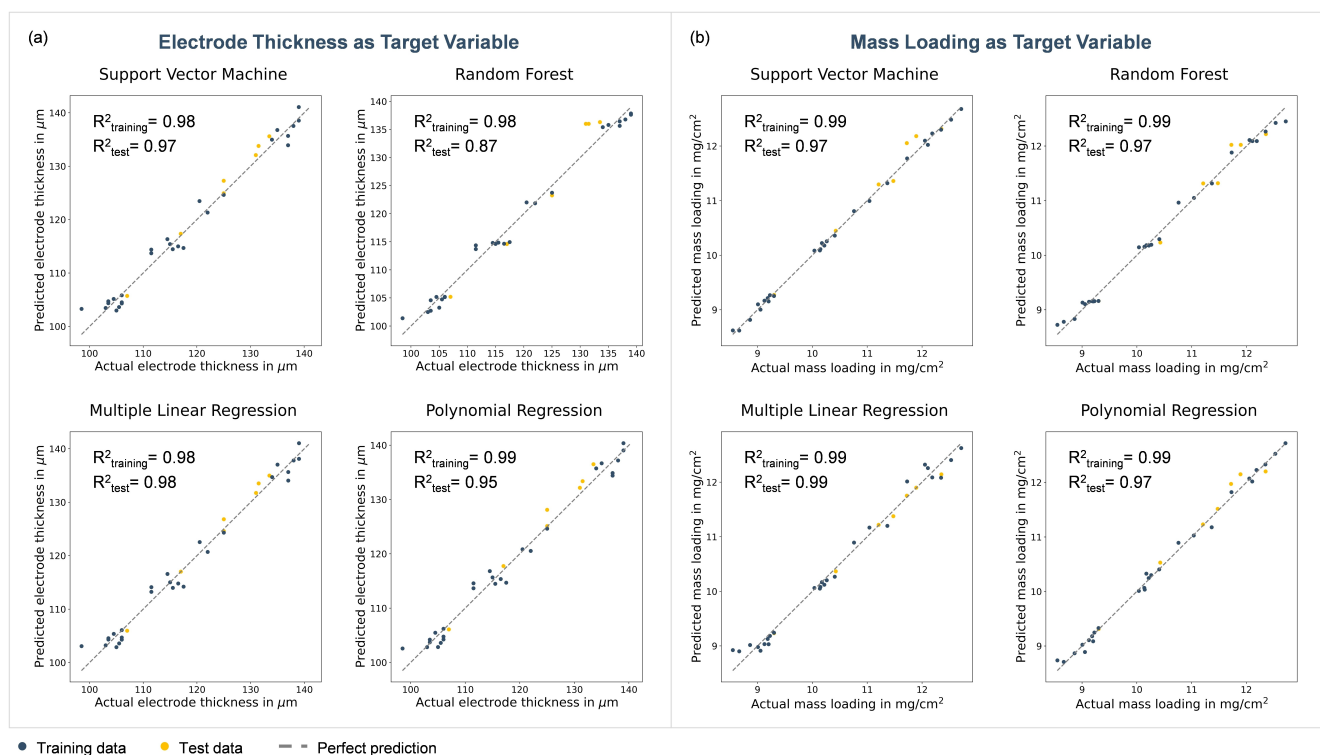
### Electrode Thickness and Mass Loading

Through continuous collection of data during the coating and drying processes, a dataset consisting of 34 data points derived from average values was obtained. This dataset was used to investigate the possibility of predicting electrode's physical properties, specifically thickness and mass loading after drying.

In case of the electrode thickness, based on the input variable of coating gap, wet mass loading and drying temperature, all four developed models showed strong efficacy, with high  $R^2$  values for both training and test dataset. The results are visualized in Figure 5(a). In case of RF, a slight overfitting can be

recognized, considering the difference between training and test performance. The correlations observed in the coating and drying processes using the doctor blade technology appears to be of a linear nature. Consequently, simpler models like MLR can effectively model the process and predict the target values.

The same input variables were used to predict the mass loading. Generally, the models showed a slightly better performance, compared to electrode thickness. This can be attributed to the impact of drying temperature on the electrode's microstructure and, subsequently, its thickness, while the mass loading remains largely unaffected. It is important to note that, when excluding the wet mass loading as input variable, the models demonstrated marginally weaker performance, with an  $R^2$  value of approximately 0.9 for the MLR. This variation can be attributed to the different measuring principles and data granularity provided by different sensors, highlighting the importance of inline measuring instruments and continuous data stream. While the coating gap is determined through two sensors located on the sides of the doctor blade, the ultrasound system utilizes a traversing setup to measure the wet mass loading area-wide. This approach contributes to greater data accuracy and granularity. The inline sensors support the characterization of the intermediate products, serving as additional quality indicators and enhancing the precision of prediction in the pursuit of achieving the predictive quality vision.



**Figure 5.** Parity plots for the developed models predicting a) electrode thickness, b) mass loading.

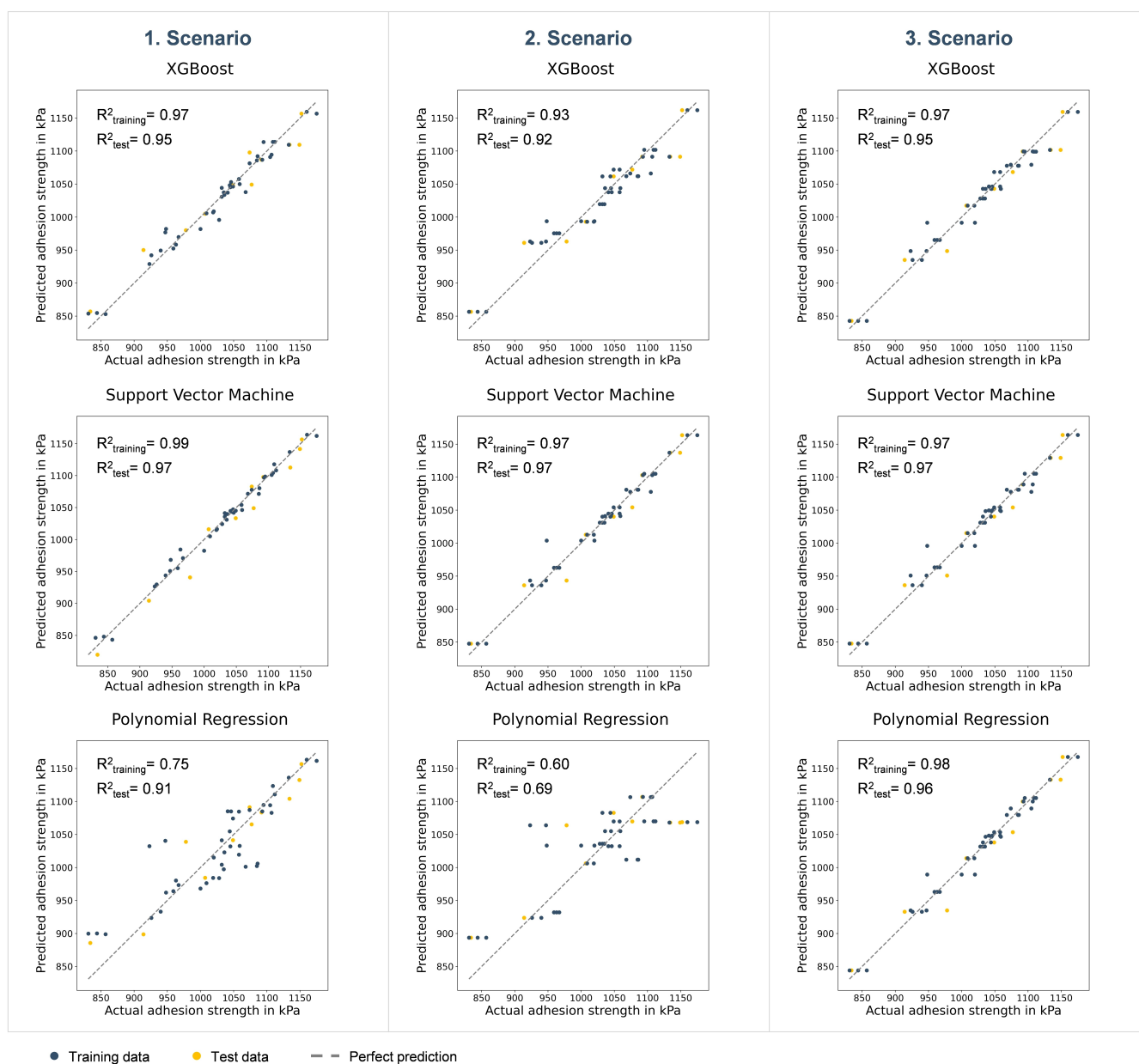


## Adhesion

The adhesion is a crucial mechanical property of the electrode that can impact the cycle performance and cell's durability.<sup>[51,52]</sup> Additionally, it plays an important role concerning the processability of the electrode throughout the production chain. From an operational perspective, various factors such as formulation,<sup>[53]</sup> mass loading,<sup>[54]</sup> drying conditions<sup>[21]</sup> and porosity<sup>[55]</sup> can impact the adhesion. According to the current state of the art,<sup>[13]</sup> adhesion can be only measured offline, hence it is also evaluated as highly complex concerning the digitalization (see Table 1). This limitation emphasizes the need for predictive solutions. The ML methods can present a promising avenue to monitor this multifaceted phenomenon without

resorting to frequent offline measurements. To explore this potential, the three scenarios, as outlined earlier, were evaluated. Among the considered modeling techniques, XGBoost, SVM, and to a certain degree, the polynomial regression demonstrated good performance. As mentioned at the beginning of this section, XGBoost was selected as an alternative to RF, to prevent overfitting and to ensure that the complex target variables can be modeled effectively. The results are presented in Figure 6.

The first scenario is based on the drying temperature and sample-specific data consisting of mass loading and porosity as input variables. In this case, XGBoost and SVM demonstrated high performance, while the polynomial regression, despite having acceptable  $R^2$  values, showed considerable scatter of the



**Figure 6.** Parity plots for the considered scenarios to predict the adhesion strength; the first scenario is based on drying temperature and sample-specific mass loading and porosity as input variables, the second scenario incorporates drying temperature, coating gap and compaction rate, and the third scenario uses a combination of process parameters and intermediate products.

data points in the parity plot. This suggests that adhesion as target variable can be modeled better using more complex models rather than simple ones. This aspect is more evident in the second scenario, in which the average values for the corresponding process parameters are adopted as input variables. The final scenario is based on a combination of process parameters and intermediate product parameters, specifically coating gap, wet mass loading, dry mass loading, compaction rate and calendering gap. Given the inherent linearity among the input variables, the polynomial regression is able to capture the variance effectively and model the adhesion, offering a performance comparable to the more complex models. It should be noted that the first scenario can be particularly of relevance, for instance, when introducing a new generation of material, for which with the help of DoE, a concise dataset can be efficiently generated. In this case, the model enhances process understanding and pinpoints the suitable parameters to be set, ensuring that the efforts invested in collecting and utilizing sample-specific data remain manageable. The third scenario envisions a continuous sensor-assisted data-driven production. Upon the development and deployment of the model, a transition in the quality management approach – particularly concerning extensive and frequent offline quality control measures – is to be expected.

Figure 7 showcases the results derived from the XML methods in the first scenario, providing insights into the analyzed interdependencies. In Figure 7(b) the findings using the SHAP method are presented, with the x-axis, displaying the SHAP values, which indicate the magnitude and direction of the feature's impact on the target variable. A higher absolute SHAP value suggests a stronger effect on the prediction. The y-axis lists the features, with the most significant one at the top. The colors serve to distinguish between different feature values. Based on the generated data, porosity is identified as the predominant parameter influencing the adhesion strength, with the low porosity leading to improvement of adhesion. This is followed by the mass loading, exhibiting a similar trend. The

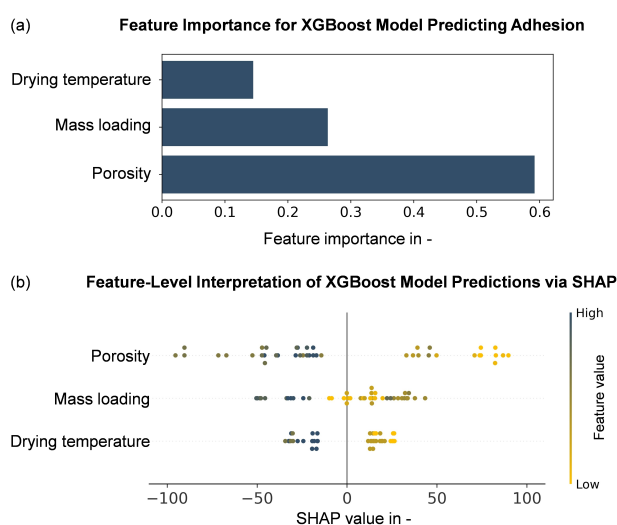
results show that a high drying temperature adversely impacts adhesion strength. The findings align with the literature, which mostly relies on isolated, full factorial, or one-factor-at-a-time analyses. Through comprehensive cross-process analysis in combination with data-driven models and XML methods, different aspects of electrode manufacturing can be pinpointed. While the feature importance plot (Figure 7a) provides an overview of the relevance of the parameters, the bee swarm plot (Figure 7b), using the SHAP values, offers both a global and granular instance-specific insight into the model's decision-making process.

### Ionic Resistance and Tortuosity

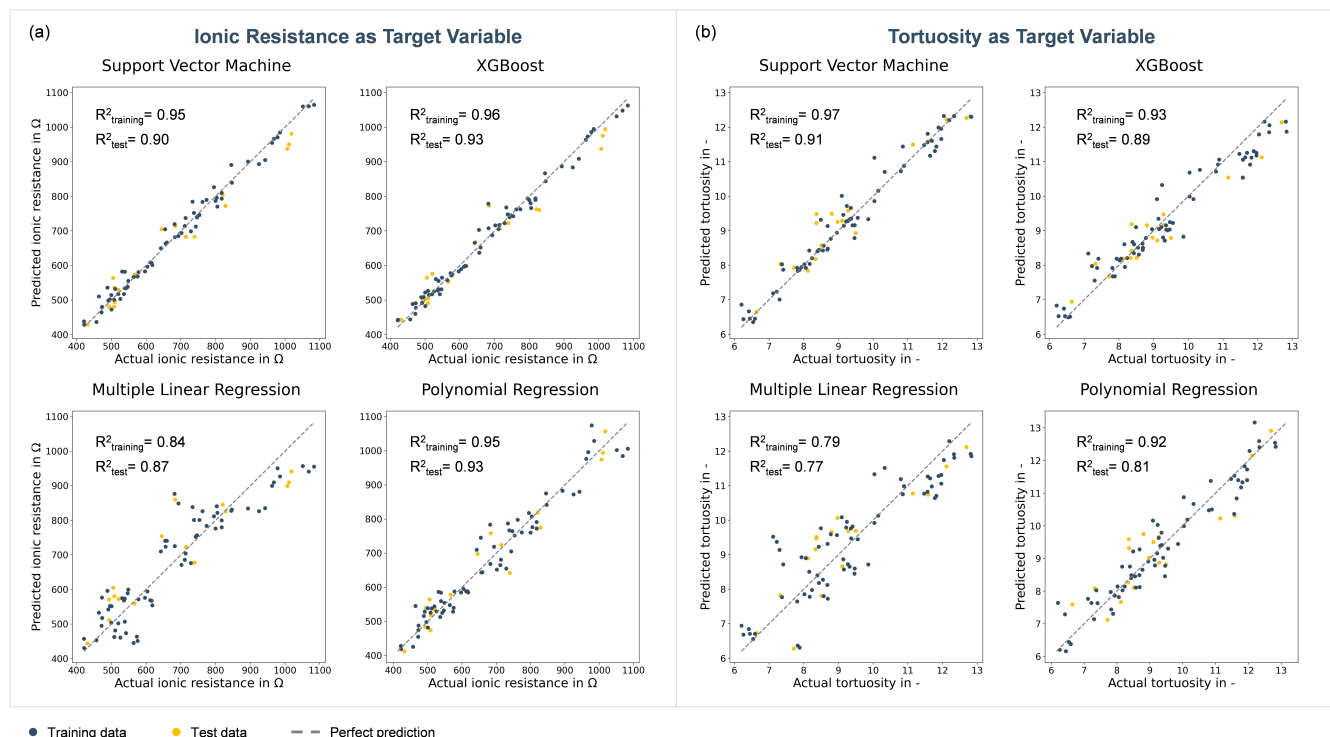
For model development based on the EIS data and following the DoE guidelines, only the data points from calendered electrodes were used. This dataset, consisting of 86 data points, was employed to predict the ionic resistance and tortuosity. Among the models developed for ionic resistance under the three defined scenarios, the MLR demonstrated the least effective performance, with an  $R^2$  value of approximately 0.7. This algorithm performed slightly weaker in predicting the tortuosity. The more complex models, XGBoost and SVM, demonstrated high performance for both target variables. The performance of the second-degree polynomial regression was found to be comparable to that of the more complex models. Figure 8 presents the results of the developed models for both ionic resistance and tortuosity in the first scenario, using the mass loading, porosity and drying temperature as input variables. In the interest of conciseness, only the models developed for the first scenario are presented in this section. The comprehensive results are available in the supplementary data.

To reveal the complex interdependencies analyzed in correlation with ionic resistance and tortuosity, the XML methods were adopted. In case of MLR models, the magnitude of the coefficients was used as an indicator representing the impact of the individual predictors on the target variable. The coefficients were normalized using Euclidean norm.<sup>[36]</sup> Figure 9 visualizes the results. Despite variations in the methods applied, the structure of the algorithms, and their accuracy, the results are mostly aligned. The only inconsistency in feature ranking was observed in the ionic resistance for the MLR model. While the XGBoost model ranked the input variables with porosity as the primary parameter influencing ionic resistance, followed by mass loading and drying temperature (cf. 9 (a)), the MLR identified mass loading as the least influential parameter (cf. 9 (c)). This discrepancy in the ranking could be attributed to the poor performance of the developed MLR model.

The feature importance (Figure 9a, b) offers insights into the significance of the input variables, while the normalized coefficient plots (Figure 9c, d) and SHAP values (Figure 9e, f) reveal the magnitude and direction of the impact of each input variable on the target value. Figure 9(c) indicates that with increased mass loading, a rise in ionic resistance is expected, which is in line with the results from the literature.



**Figure 7.** Results of the XML methods revealing features' relevance and their impact on adhesion strength using a) MDI and b) SHAP values.



**Figure 8.** Parity plots predicting a) ionic resistance, b) tortuosity based on electrode's mass loading, drying temperature and porosity.

Contrary to expectations, the results suggested that tortuosity is impacted by mass loading. Specifically, as depicted in Figure 9(d and f), high mass loading adversely affected the tortuosity. This can potentially be attributed to the cumulative effects of high mass loading and temperature on electrode thickness, and consequently the tortuosity, as outlined in Equation (2).

As an additional verification step for the adopted methodology and the developed data-driven models, the obtained results were compared with tortuosity estimations based on the Bruggeman relationship.<sup>[56]</sup> According to Bruggeman, for spherical particles, there is an interdependency between porosity and tortuosity, following the equation below, with  $\alpha = 0.5$ :<sup>[27,57]</sup>

$$\tau = \varepsilon^{-\alpha} \quad (4)$$

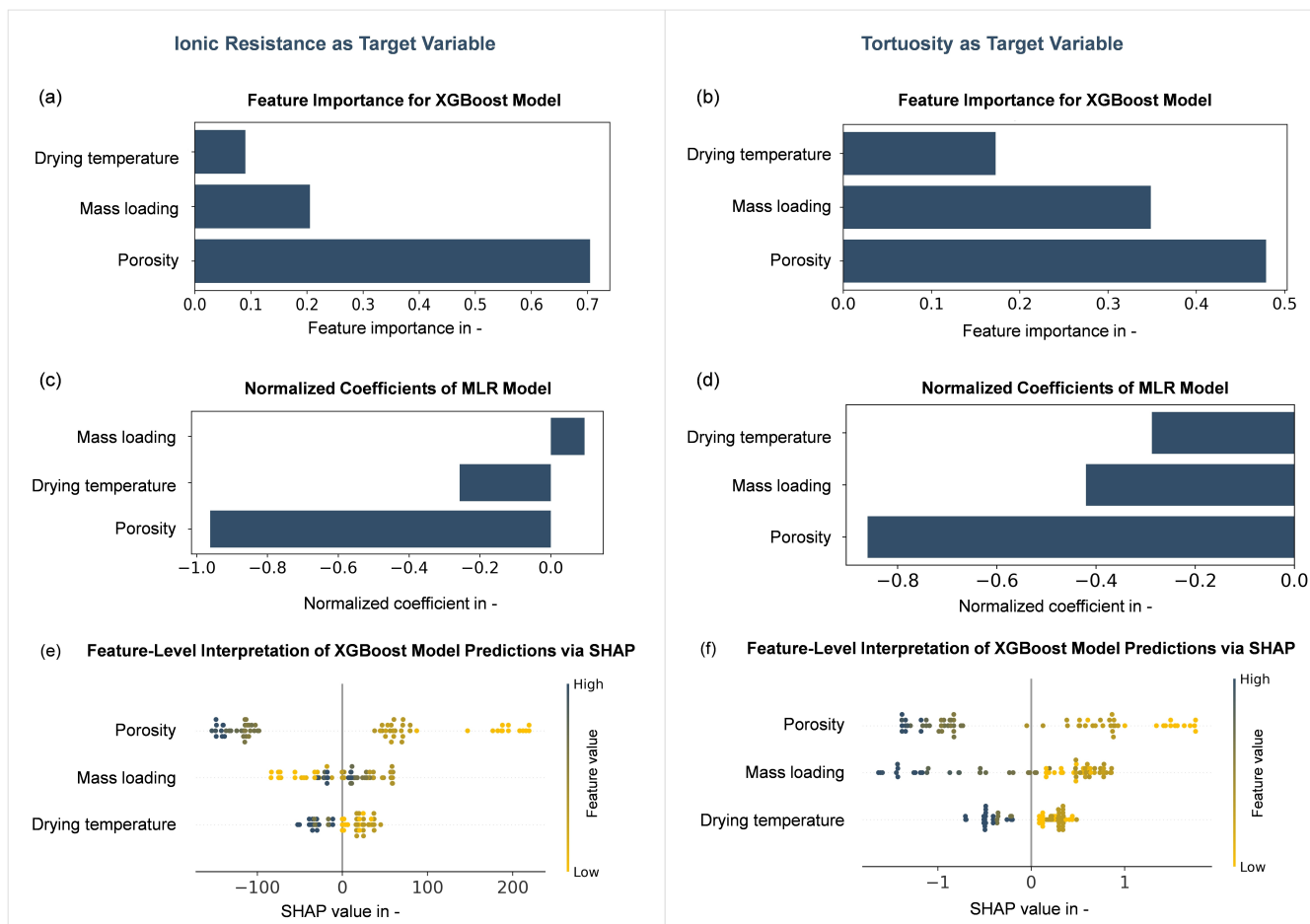
However, this simple assumption does not adequately capture the complexities of electrode manufacturing and the diverse materials used with different particle sizes and shapes.<sup>[52,57]</sup> MAYILVAHANAN ET AL.<sup>[58]</sup> introduced a modification to the Bruggeman relation based on a pseudo-two-dimensional (P2D) model, applied to NCM electrodes. The adjusted Bruggeman relation was reported as:

$$\tau = 2.3 \varepsilon^{-0.31} \quad (5)$$

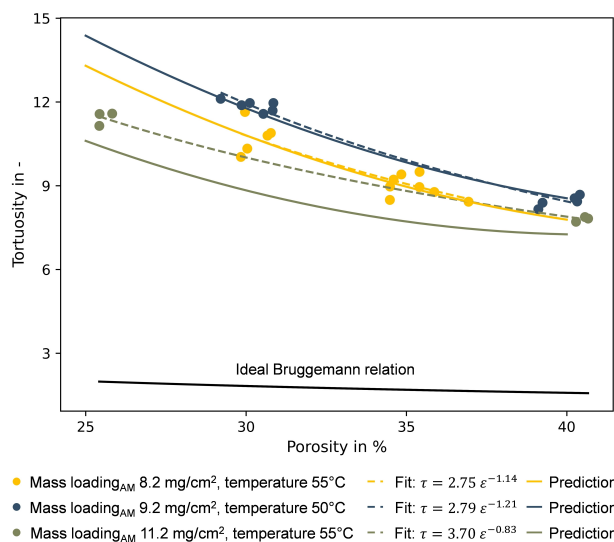
Building upon these findings, six configurations from the existing dataset were identified and grouped into three categories. Each category featured configurations with similar mass loading and drying temperature but two different

porosities. Based on the variations in the porosity, it was possible to fit the adjusted Bruggeman relation for each category, following Equation (5). Figure 10 illustrates the data points, the corresponding fits according to the Bruggeman relation and the curve generated using the developed polynomial model. There is a strong agreement between the data points, the fitted Bruggeman curve, and predicted tortuosity values, particularly for two of the three mass loadings. To understand the acceptable, yet consistent discrepancy observed for active mass loading of  $11.2 \text{ mg/cm}^2$ , the Fraction of Design Space (FDS) plot was adopted.<sup>[59]</sup> The FDS plot offers a comprehensive overview of the range of values and distribution of prediction variance throughout the entire analyzed experimental space.<sup>[60]</sup> Figure 11 visualizes the smoothed FDS plot exemplary for the two developed polynomial and SVM models, predicting the tortuosity. The plot reveals that both models exhibit a relatively low variance across the majority of the design space, an advantage of the I-optimal design over other DoE methods.<sup>[18]</sup> However, approximately 10% of the design space displays a variance higher than 1, with the polynomial model exhibiting a larger deviation than the SVM model.

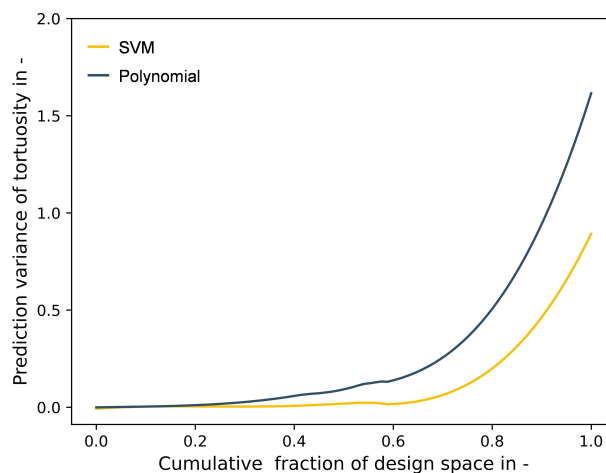
The FDS can be used to pinpoint the areas within the design space that may necessitate refinement or additional data points to enhance the model's performance. Ultimately, by considering the results from comparing the modified Bruggeman curve with the polynomial model, along with the model's performance, the portion of the design space, and the significance of the variance, the robustness and efficacy of the adopted methodology and the developed models can be confirmed.



**Figure 9.** Results of the XML methods and MLR coefficients revealing features' relevance and their impact on the prediction of ionic resistance (left) and tortuosity (right).



**Figure 10.** Tortuosity versus porosity plot for six configurations, categorized by mass loading and drying temperature. The data points are represented by dots, the fitted curves using the modified Bruggemann relation are shown in dashed lines, and the predictions using the developed polynomial models are illustrated with solid lines.



**Figure 11.** Exemplary FDS plot for two of the developed models predicting the tortuosity.

### Discharge Capacities at Different C-Rates

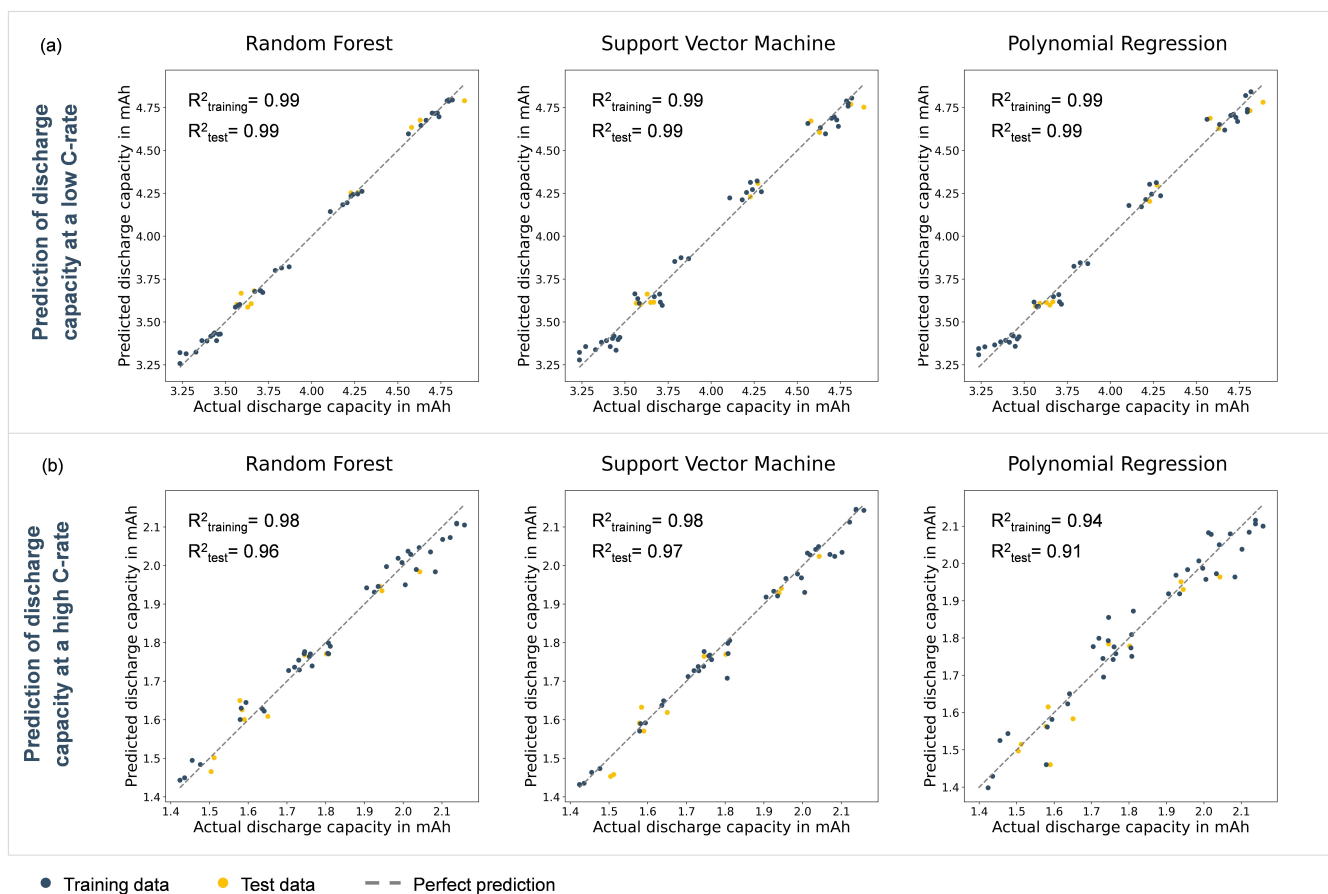
To assess the impact of the electrode manufacturing parameters on the cell performance, the discharge capacities at four

distinct C-rates, representing a range of applications, were investigated.

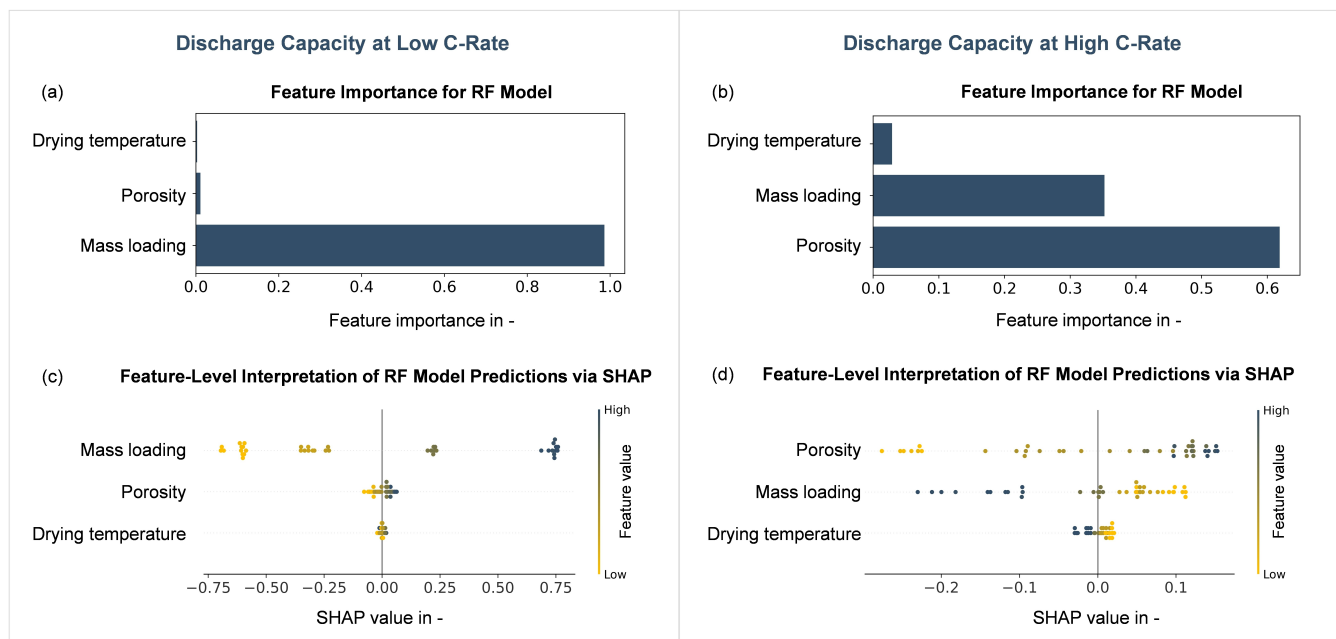
This section presents the results for the discharge capacity at 0.1 C (hereafter referred to as low C-rate) and 5 C, termed the high C-rate. Since the findings for 1 C and 3 C were comparable to the ones included here, the associated results are provided only in the supplementary data. Among the considered algorithms, RF, SVM, and polynomial regression demonstrated high performance. The MLR was able to predict the discharge capacity at low C-rates but struggled to maintain satisfactory performance at higher rates. This indicates that the cell behavior at higher C-rates is influenced by more intricate factors that cannot be captured using a simple linear regression or a limited number of data points. Concerning the three scenarios defined, the results aligned with those discussed for the adhesion. The cell-specific input variables were found to improve the prediction accuracy of the models, while the combination of process and intermediate parameters demonstrated comparable results. The details can be found in the supplementary data. Figure 12 illustrates the results for the first scenario. All models exhibit high performance at low C-rate. Given the underlying principles of the lithium-ion battery, it is evident that the cell capacity at a low C-rate is dominated by the mass loading and hence the models were able to capture this linear relation effectively. At higher C-rates, the cell

performance is impacted by an interplay of factors including mass loading,<sup>[61]</sup> porosity,<sup>[62]</sup> and drying temperature.<sup>[63]</sup> Nonetheless, the models exhibited commendable performance.

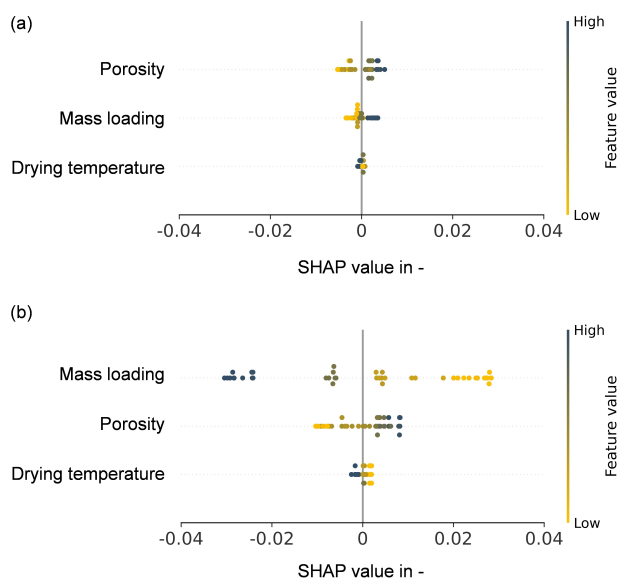
Figure 13 reveals the significance of the features and their impact on the discharge cell capacity at different C-rates. Both adopted methods, Feature Importance using MDI and SHAP value, confirmed that the discharge capacity at a low C-rate was consistently influenced by the mass loading (Figure 12a and c). At a high C-rate, within the analyzed range, porosity was identified as the dominant factor, with low porosity resulting in a lower discharge capacity. This was followed by mass loading, showing an inverse effect. A similar trend was observed with the drying temperature. Additionally, the gravimetric cell capacity was analyzed as a target variable, given its significance in electrode design for high-energy batteries.<sup>[64]</sup> At a low C-rate the models demonstrated a moderate performance (see Table 5). At a high C-rate a comparable performance to the results of the discharge capacities was observed. Figure 14 highlights the relevance of the analyzed parameters in relation to the gravimetric discharge capacity at low and high C-rates. As expected, at low C-rate, the mass loading does not have a pronounced influence. With the porosity marked as the most relevant parameter. In contrast, at high C-rate, the impact of mass loading became more evident, followed by the porosity and drying temperature.



**Figure 12.** Parity plots of the developed models predicting the discharge capacity at a) 0.1 C, b) 5 C based on the drying temperature and cell-specific mass loading and porosity.



**Figure 13.** Results of the XML methods highlighting the relevance of the features with respect to the cell discharge capacity at low (left) and high (right) C-rates.



**Figure 14.** SHAP values revealing the impact of parameters on the gravimetric discharge capacity at a) 0.1 C, b) 5 C.

The models developed in this study, including the input and target variables and the evaluation metrics, are summarized in Table 5. Additionally, the supplementary data includes the results for the models developed using only process parameters, as well as those developed using a combination of both product and process parameters.

## Conclusions

This article presented a systematic efficient analysis of the interdependencies in electrode manufacturing. Through the integration of DoE and data-driven models, in combination with the findings from the previous study,<sup>[12]</sup> a comprehensive cross-process analysis was conducted. The study centered on three quality-relevant parameters from coating, drying and calendaring processes, assessing their impact on both the mechanical and electrochemical characteristics of the electrodes. Several research gaps<sup>[6]</sup> were addressed: (I) the drying process, recognized as an energy-intensive and multifaceted step, was analyzed in combination with its closely interrelated processes and the relevant parameters; (II) the framework was demonstrated on the graphite anode, which has received limited attention in the data-driven literature; (III) less explored target variables, such as adhesion and tortuosity, were investigated; (IV) aiming for a holistic framework, the study offered a detailed level of information, covering various aspects from identification of relevant parameters and data generation to data preprocessing and modeling. The showcased framework can be used as a blueprint for efficient generation of a dataset, particularly when introducing new materials or for undertaking specific optimization studies. The former can be considered as a valuable contribution to the material acceleration platforms.<sup>[65]</sup> Data-driven models, in combination with the XML methods, can help to unravel the complex interdependencies in electrode manufacturing. The findings facilitate the process understanding and aid in the identification of suitable process parameters.

**Table 5.** Summary of the developed models, including the input and target variables and the evaluation metrics.

Target variable	Input variables	Model	Evaluation metrics				
			$R^2_{\text{training}}$	$R^2_{\text{test}}$	RMSE <sub>training</sub>	RMSE <sub>test</sub>	MSE <sub>training</sub> /MSE <sub>test</sub>
Mass loading	Coating gap Wet mass loading Drying temperature	MLR	0.99	0.99	0.2	0.1	2.8
		RF	0.99	0.97	0.1	0.2	0.3
		Polynomial	0.99	0.97	0.2	0.1	0.3
		SVM	0.99	0.97	0.1	0.2	0.1
Electrode thickness	Coating gap Wet mass loading Drying temperature	MLR	0.98	0.98	2	1	2.1
		RF	0.98	0.87	1	3	0.2
		Polynomial	0.98	0.95	2	2	0.9
		SVM	0.98	0.97	2	2	1.4
Mass loading	Coating gap Drying temperature	MLR	0.88	0.91	0.5	0.3	2.5
		RF	0.99	0.80	0.1	0.4	0.1
		Polynomial	0.89	0.97	0.4	0.2	6.4
		SVM	0.98	0.80	0.2	0.4	0.2
Electrode thickness	Coating gap Drying temperature	MLR	0.83	0.87	5	3	3.0
		RF	0.97	0.74	2	4	0.3
		Polynomial	0.84	0.89	5	3	3.3
		SVM	0.96	0.56	3	6	0.2
Adhesion strength	Mass loading Drying temperature Porosity	XGBoost	0.97	0.95	14	22	0.4
		Polynomial	0.75	0.91	40	28	1.9
		SVM	0.99	0.97	8	18	0.2
Adhesion strength	Coating gap Drying temperature Compaction rate	XGBoost	0.93	0.91	20	28	0.5
		Polynomial	0.60	0.69	50	53	0.9
		SVM	0.97	0.97	13	16	0.6
Adhesion strength	Coating gap Wet mass loading Drying temperature Dry mass loading Compaction rate Calendering gap	XGBoost	0.97	0.95	14	21	0.4
		Polynomial	0.98	0.96	11	19	0.4
		SVM	0.97	0.97	13	16	0.7
Ionic resistance	Mass loading Drying temperature Porosity	XGBoost	0.96	0.93	35	55	0.4
		MLR	0.84	0.87	68	71	0.9
		Polynomial	0.95	0.93	39	52	0.6
		SVM	0.95	0.90	39	63	0.4
Tortuosity	Mass loading Drying temperature Porosity	XGBoost	0.93	0.89	0.5	0.5	0.8
		MLR	0.79	0.77	0.8	0.8	1.2
		Polynomial	0.92	0.81	0.5	0.7	0.5
		SVM	0.97	0.91	0.3	0.5	0.5
Discharge capacity at 0.1 C	Mass loading Drying temperature Porosity	RF	0.99	0.99	0.0277	0.0494	0.3
		Polynomial	0.99	0.99	0.0549	0.0575	0.9
		SVM	0.99	0.99	0.0605	0.0582	1.1
Discharge capacity at 5 C	Mass loading Drying temperature Porosity	RF	0.98	0.96	0.0331	0.0374	0.8
		Polynomial	0.94	0.91	0.0509	0.0532	0.9
		SVM	0.98	0.97	0.0275	0.0320	0.7
Gravimetric discharge capacity at 0.1 C	Mass loading Drying temperature Porosity	RF	0.84	0.77	0.0020	0.0032	0.4
		Polynomial	0.56	0.77	0.0034	0.0032	1.1
		SVM	0.78	0.65	0.0024	0.0040	0.4

Table 5. continued							
Target variable	Input variables	Model	Evaluation metrics				
			$R^2_{\text{training}}$	$R^2_{\text{test}}$	RMSE <sub>training</sub>	RMSE <sub>test</sub>	MSE <sub>training</sub> /MSE <sub>test</sub>
Gravimetric discharge capacity at 5 C	Mass loading	RF	0.99	0.98	0.0023	0.0021	1.2
	Drying temperature	Polynomial	0.98	0.96	0.0033	0.0032	1.0
	Porosity	SVM	0.99	0.99	0.0016	0.0018	0.8

The dataset for the mechanical and electrochemical characteristics using full coin cells consisted of 51 data points for the 17 different electrode configurations produced using I-optimal DoE method. For the EIS, data was collected from 86 symmetric cells. To the best of the authors' knowledge, this dataset is the most comprehensive collection of impedance spectroscopy data available to date, reflecting various aspects of electrode manufacturing from electrodes produced on a pilot line. The comprehensive data can be found in the supplementary section.

From the modeling perspective, the study covered a range of techniques, from more complex models such as RF, XGBoost, and SVM to simpler methods such as MLR and polynomial regression. The capability of different models in predicting various target variables was demonstrated. For some applications, such as mass loading and electrode thickness, a simple algorithm proved to be sufficient. However, when addressing intricate and multifaceted targets, such as adhesion, the need for more sophisticated models became evident. Given the structure of the study, which was based on a train-test split, XGBoost was chosen over RF for certain targets, specifically adhesion, ionic resistance, and tortuosity. This decision was driven by a couple of considerations. The XGBoost offers built-in regularization, which inherently combats overfitting, ensuring that the model does not become too closely tailored to the training data. Additionally, it is able to capture and represent the complex interdependencies existing in the analyzed system. To address the need for explainability of complex models and assist with the process understanding, XML methods were integrated. These methods not only provide clarity on the decision-making mechanism of the models but also offer valuable insights into the complex interdependencies and influential factors. Through XML methods, a more transparent and informed decision-making process can be achieved, fostering greater trust of domain experts in ML and accelerating its practical applications in real-world scenarios.

The significance of inline sensor data for the characterization of intermediate parameters was demonstrated through the consideration of three scenarios for the development of models. It was observed that the models built exclusively on the process parameters yielded the least effective results. In contrast, when combining process parameters with intermediate products, the models' performance was almost equivalent to those utilizing sample-specific product parameters.

It should be noted that the derived insights should be contextualized within the boundaries of the examined ranges. For instance, exploring a larger range for mass loading might lead to a different ranking when it comes to adhesion strength. With respect to adhesion, as detailed in the data preprocessing and analysis section, in the case of the low mass loading, except for a local optimum at the low temperature, no significant impact of higher drying temperatures was observed. To gain a deeper understanding of the drying process, further analysis of the drying temperatures, for instance in the first dryer, coupled with the drying speed, would be beneficial. Hence, future work will focus on expanding the dataset and broadening the scope of the parameters analyzed.

## Supporting Information

Supporting information is available online.

## Nomenclature

AM	Active Material
DoE	Design of Experiments
EIS	Electrochemical Impedance Spectroscopy
FDS	Fraction of Design Space
MDI	Mean Decrease Impurity
ML	Machine Learning
MLR	Multiple Linear Regression
MSE	Mean Squared Error
NCM	Nickel Cobalt Magnesium
P2D	Pseudo-Two-Dimensional
RF	Random Forest
RMSE	Root Mean Squared Error
RSM	Response Surface Methodology
SHAP	SHapley Additive exPlanations
SVM	Support Vector Machine
TLM	Transmission Line Model
VIF	Variance Inflation Factor
XGBoost	eXtreme Gradient Boosting
XML	eXplainable Machine Learning

## Author Contributions

**Sajedeh Haghi:** Conceptualization, Methodology, Investigation, Software, Formal analysis, Data curation, Validation, Visual-



ization, Writing – Original Draft. **Josef Keilhofer**: Formal analysis, Writing – Review & Editing. **Nico Schwarz**: Software, Data Curation. **Pengdan He**: Investigation, Data Curation. **Rüdiger Daub**: Funding acquisition, Resources, Supervision, Writing – Review & Editing. “All authors have read and agreed to the published version of the manuscript.”

## Acknowledgements

This research work was conducted as part of the TrackBatt project (grant number: 03XP0310), funded by the German Federal Ministry of Education and Research, abbreviated BMBF. The authors gratefully acknowledge the financial support provided by BMBF. The authors would also like to thank Andreas Mayr, Filip Dorau and Roman Mazur for their assistance in electrode manufacturing and coin cell assembly. Open Access funding enabled and organized by Projekt DEAL.

## Conflict of Interests

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

**Keywords:** tortuosity · adhesion · lithium-ion battery · quality · DoE

- [1] J. Li, J. Fleetwood, W. B. Hawley, W. Kays, *Chem. Rev.* **2022**, *122*, 903–956.
- [2] A. Kwade, W. Haselrieder, R. Leithoff, A. Modlinger, F. Dietrich, K. Droeder, *Nat. Energy* **2018**, *3*, 290–300.
- [3] S. Haghi, H.-C. Töpfer, F. J. Günter, G. Reinhart, *Procedia CIRP* **2021**, *104*, 1155–1160.
- [4] T. Günther, N. Billot, J. Schuster, J. Schnell, F. B. Spingler, H. A. Gasteiger, *AM + R* **2016**, *1140*, 304–311.
- [5] T. Lombardo, M. Duquesnoy, H. El-Bouysidy, F. Årén, A. Gallo-Bueno, P. B. Jørgensen, A. Bhowmik, A. Demortière, E. Ayerbe, F. Alcaide et al., *Chem. Rev.* **2022**, *122*, 10899–10969.
- [6] S. Haghi, M. F. V. Hidalgo, M. F. Niri, R. Daub, J. Marco, *Batteries & Supercaps* **2023**, *6*, e202300046.
- [7] A. Turetsky, S. Thiede, M. Thomitzek, N. von Drachenfels, T. Pape, C. Herrmann, *Energy Technol.* **2020**, *8*, 1900136.
- [8] M. F. Niri, K. Liu, G. Apachitei, L. A. Román-Ramírez, M. Lain, D. Widanage, J. Marco, *Energy AI* **2022**, *7*, 100129.
- [9] R. P. Cunha, T. Lombardo, E. N. Primo, A. A. Franco, *Batteries & Supercaps* **2020**, *3*, 60–67.
- [10] K. Liu, M. F. Niri, G. Apachitei, M. Lain, D. Greenwood, J. Marco, *Control Engineering Practice* **2022**, *124*, 105202.
- [11] S. X. Drakopoulos, A. Gholamipour-Shirazi, P. MacDonald, R. C. Parini, C. D. Reynolds, D. L. Burnett, B. Pye, K. B. O'Regan, G. Wang, T. M. Whitehead et al., *Cell Rep. Phys. Sci.* **2021**, *2*, 100683.
- [12] S. Haghi, A. Summer, P. Bauerschmidt, R. Daub, *Energy Technol.* **2022**, *10*, 2200657.
- [13] S. Haghi, M. Leeb, A. Molzberger, R. Daub, *Energy Technol.* **2023**, *11(9)*, 2300364.
- [14] R. Arboretti, R. Ceccato, L. Pegoraro, L. Salmaso, *Qual. Reliab. Eng.* **2022**, *38*, 1131–1156.
- [15] L. A. Román-Ramírez, J. Marco, *Appl. Energy* **2022**, *320*, 119305.
- [16] R. H. Myers, D. C. Montgomery, C. M. Anderson-Cook, *Response surface methodology. Process and Product Optimization Using Designed Experiments*, John Wiley & Sons, Hoboken, New Jersey, **2016**.
- [17] D. C. Montgomery, *Design and analysis of experiments*, John Wiley & Sons, Hoboken, New Jersey, **2020**.
- [18] B. Jones, P. Goos, *J. Qual. Technol.* **2012**, *44*, 85–101.
- [19] A. Du Baret de Limé, T. Lein, S. Maletti, K. Schmal, S. Reuber, C. Heubner, A. Michaelis, *Batteries & Supercaps* **2022**, *5*, e202200239.
- [20] T. Günther, D. Schreiner, A. Metkar, C. Meyer, A. Kwade, G. Reinhart, *Energy Technol.* **2020**, *8*, 1900026.
- [21] S. Jaiser, N. Sanchez Salach, M. Baunach, P. Scharfer, W. Schabel, *Drying Technol.* **2017**, *35*, 1807–1817.
- [22] S. Jaiser, A. Friske, M. Baunach, P. Scharfer, W. Schabel, *Drying Technol.* **2017**, *35*, 1266–1275.
- [23] M. H. Kutner, C. Nachtsheim, J. Neter, L. William, *Applied linear statistical models*, McGraw-Hill Irwin, Boston, Mass., **2005**.
- [24] L. Hille, H.-C. Toepfer, C. Schriever, J. Keilhofer, M. P. Noecker, M. F. Zaeh, *J. Electrochem. Soc.* **2022**, *169*, 60518.
- [25] W. Haselrieder, B. Westphal, H. Bockholt, A. Diener, S. Höft, A. Kwade, *Int. J. Adhes. Adhes.* **2015**, *60*, 1–8.
- [26] N. Ogihara, S. Kawachi, C. Okuda, Y. Itou, Y. Takeuchi, Y. Ukyo, *J. Electrochem. Soc.* **2012**, *159*, A1034–A1039.
- [27] J. Landesfeind, J. Hattendorff, A. Ehrl, W. A. Wall, H. A. Gasteiger, *J. Electrochem. Soc.* **2016**, *163*, A1373–A1387.
- [28] M. Murbach, B. Gerwe, N. Dawson-Elli, L. Tsui, *JOSS* **2020**, *5*, 2349.
- [29] V. F. Lvovich, *Impedance spectroscopy. Applications to electrochemical and dielectric phenomena*, John Wiley & Sons, Hoboken, New Jersey, **2012**.
- [30] A. van Bommel, R. Divigalpitiya, *J. Electrochem. Soc.* **2012**, *159*, A1791–A1795.
- [31] R. Morasch, J. Landesfeind, B. Suthar, H. A. Gasteiger, *J. Electrochem. Soc.* **2018**, *165*, A3459–A3467.
- [32] J. Kumberg, M. Baunach, J. C. Eser, A. Altvater, P. Scharfer, W. Schabel, *Energy Technol.* **2021**, *9*, 2100549.
- [33] S. J. Russell, P. Norvig, *Artificial intelligence. A modern approach*, Prentice-Hall, Upper Saddle River, NJ, **2010**.
- [34] K. P. Murphy, *Machine learning. A probabilistic perspective*, MIT Press, **2012**.
- [35] G. James, D. Witten, T. Hastie, R. Tibshirani, J. Taylor, *An Introduction to Statistical Learning. With Applications in R*, Springer, New York, NY, **2013**.
- [36] T. Hastie, R. Tibshirani, J. H. Friedman, *The elements of statistical learning. Data mining, inference, and prediction*, Springer, New York, NY, **2017**.
- [37] R. Roscher, B. Bohn, M. F. Duarte, J. Garcke, *IEEE Access* **2020**, *8*, 42200–42216.
- [38] S. Wächter, B. Mittelstadt, C. Russell, *JOLT* **2018**, *31*, 841, DOI: <https://doi.org/10.2139/ssrn.3063289>.
- [39] A. Rai, *JAMS* **2020**, *48*, 137–141.
- [40] M. Faraji Niri, K. Aslansefat, S. Haghi, M. Hashemian, R. Daub, J. Marco, *Energies* **2023**, *16*, 6360.
- [41] L. Breiman, *Mach. Learn.* **2001**, *45*, 5–32.
- [42] L. Breiman, *Manual on setting up, using, and understanding random forests*, Statistics Department University of California Berkeley, CA, USA, **2002**.
- [43] G. Louppe, L. Wehenkel, A. Sutera, P. Geurts, *Adv. Neural Inf. Process.* **2013**, *26*, 431–439.
- [44] S. Nembrini, I. R. König, M. N. Wright, *Bioinformatics* **2018**, *34*, 3711–3718.
- [45] S. M. Lundberg, S.-I. Lee, *arXiv:1705.07874* **2017**, *30*, DOI: <https://doi.org/10.48550/arXiv.1705.07874>.
- [46] L. Antwarg, R. M. Miller, B. Shapira, L. Rokach, *Expert Syst. Appl.* **2021**, *186*, 115736.
- [47] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, **2000**.
- [48] D. N. Moriasi, J. G. Arnold, M. W. van Liew, R. L. Bingner, R. D. Harmel, T. L. Veith, *Trans ASABE* **2007**, *50(3)*, 885–900.
- [49] W. Haselrieder, S. Ivanov, D. K. Christen, H. Bockholt, A. Kwade, *ECS Trans.* **2013**, *50*, 59–70.
- [50] B. Ratner, *J. Target Meas Anal Mark* **2009**, *17*, 139–142.
- [51] H. Zheng, L. Zhang, G. Liu, X. Song, V. S. Battaglia, *J. Power Sources* **2012**, *217*, 530–537.
- [52] W. B. Hawley, J. Li, *J. Energy Storage* **2019**, *25*, 100862.

- [53] N. Billot, T. Günther, D. Schreiner, R. Stahl, J. Kranner, M. Beyer, G. Reinhart, *Energy Technol.* **2020**, *8*, 1801136.
- [54] B. G. Westphal, A. Kwade, *J. Energy Storage* **2018**, *18*, 509–517.
- [55] C. Sangrós Giménez, C. Schilde, L. Froböse, S. Ivanov, A. Kwade, *Energy Technol.* **2020**, *8*, 1900180.
- [56] D. A. G. Bruggeman, *Ann. Phys.* **1935**, *416*, 636–664.
- [57] B. Tjaden, S. J. Cooper, D. J. L. Brett, D. Kramer, P. R. Shearing, *Curr. Opin. Chem. Eng.* **2016**, *12*, 44–51.
- [58] K. S. Mayilvahanan, Z. Hui, K. Hu, J. Kuang, A. H. McCarthy, J. Bernard, L. Wang, K. J. Takeuchi, A. C. Marschilok, E. S. Takeuchi et al., *J. Electrochem. Soc.* **2021**, *168*, 70537.
- [59] A. Zahran, C. M. Anderson-Cook, R. H. Myers, *JQT* **2003**, *35*, 377–386.
- [60] A. Ozol-Godfrey, C. Anderson-Cook, T. J. Robinson, *J. Stat. Plan. Inference* **2008**, *138*, 203–219.
- [61] D. Parikh, T. Christensen, J. Li, *J. Power Sources* **2020**, *474*, 228601.
- [62] M. Singh, J. Kaiser, H. Hahn, *Batteries* **2016**, *2*, 35.
- [63] S. Jaiser, M. Müller, M. Baunach, W. Bauer, P. Scharfer, W. Schabel, *J. Power Sources* **2016**, *318*, 210–219.
- [64] C. Heubner, A. Nickol, J. Seeba, S. Reuber, N. Junker, M. Wolter, M. Schneider, A. Michaelis, *J. Power Sources* **2019**, *419*, 119–126.
- [65] M. M. Flores-Leonar, L. M. Mejía-Mendoza, A. Aguilar-Granda, B. Sanchez-Lengeling, H. Tribukait, C. Amador-Bedolla, A. Aspuru-Guzik, *Curr. Opin. Green Sustain. Chem.* **2020**, *25*, 100370.

---

Manuscript received: October 11, 2023

Revised manuscript received: November 23, 2023

Accepted manuscript online: November 27, 2023

Version of record online: December 7, 2023