industries

MDPI

Article

Interactive Visual Data Analysis for Quality Optimization in Aluminum Direct Chill Casting

Manuela Larissa Schreyer [1,*], Alexander Gerber [2], Steffen Neubert [2] and Wolfgang Trutschnig [2,3]

[1] Austria Metall GmbH, Lamprechtshausener Str. 61, 5282 Ranshofen, Austria
[2] AMAG Casting GmbH, Lamprechtshausener Str. 61, 5282 Ranshofen, Austria; alexander.gerber@amag.at (A.G.); steffen.neubert@amag.at (S.N.); wolfgang.trutschnig@plus.ac.at (W.T.)
[3] Department for Artificial Intelligence & Human Interfaces, University of Salzburg, Hellbrunnerstrasse 34, 5020 Salzburg, Austria
* Correspondence: manuela.schreyer@amag.at

Abstract

In view of the increasing digitalization of industry, data science and data visualization are assuming an increasingly pivotal role in numerous industry sectors, facilitating enhanced monitoring and analysis of production processes. This development is primarily driven by growing demands for product quality and sustainability. In a multitude of industrial contexts, large volumes of high-resolution sensor data are continuously recorded throughout production. To make this huge amount of data usable for technologists, an interactive visual analytics tool for the analysis of production data has been developed. The tool enables users to explore complex process data, identify relevant patterns, and derive data-driven insights for quality optimization. This, in turn, leads to reduced scrap rates, as well as improved plant utilization and delivery performance. The present paper demonstrates the effectiveness of the proposed approach using a selected industrial use case and highlights key insights gained from its application in practice.

Keywords: process data visualization; quality optimization; aluminum casting; discretization; data science; big data

1. Introduction

In view of the increasing digitalization of industry and the growing demands for product quality and sustainability [1–3], this article examines continuous production time-series data recorded in industrial processes and presents the application of an interactive visual analytics tool that enables technologists to explore complex production data, identify relevant process patterns, and derive data-driven insights for quality optimization.

As an industrial use case, the study focuses on the casting of aluminum ingots at AMAG Austria Metall AG, where large volumes of high-resolution sensor data are generated during production. Figure 1 shows system operators during the casting process at AMAG.

After casting, the ingots are rolled into sheets and plates. A schematic representation of the production process is provided in Figure 2. The process begins with the input material, which is primarily aluminum scrap. The input material is then melted over a period of several hours and cast into ingots. Each batch consists of up to six ingots, with an approximate length of seven meters and a width of 0.5 m. Subsequently, the ingots are rolled into sheets and plates [4]. Depending on the customers' requirements, it is essential that high-quality standards are met [5]. This is confirmed through a final

Check for updates

Academic Editor: Ying Liu

Received: 15 June 2026
Revised: 22 July 2026
Accepted: 16 September 2026
Published: 4 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

laboratory test. The finished products are then utilized in various industrial sectors, with major applications including aerospace, automotive, sports, and packaging industries, among numerous others [1–3,6].



Figure 1. System operator during the casting process.

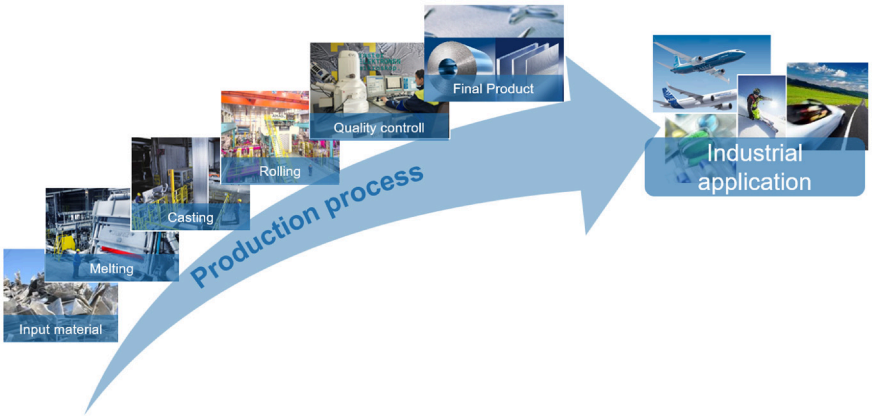


Figure 2. Schematic and simplified illustration of the production process.

During the production of an aluminum ingot in the foundry, a multitude of sensor readings are measured over a period of several hours, a situation that is characteristic of many industrial production processes. Due to the substantial volume of high-resolution data recorded over the entire production period, these measurements cannot be effectively analyzed without suitable data transformation and visual analytics approaches [7].

The ongoing digital transformation of the industry is driving the need for advanced data analysis and visualization tools, facilitating continuous improvement and monitoring of production processes [1–3,6,8]. At the same time, reducing rejection rates, improving plant utilization, and shortening delivery times remain key objectives in industrial practice.

To address this need, the visual analysis tool ("ADAM Process Data", an acronym for aluminum production Data Analysis and Monitoring for Process Data) is developed in collaboration with product technologists, enabling the visual analysis of large volumes of process data. Previous work has also presented a prototype for the visual analysis of product quality data at AMAG [9,10], which will be referenced in the following sections.

While the underlying methodology is described in detail in previous work (see, for example, [3]) and will be referenced in the respective sections, this article focuses on its application within an interactive visual analytics environment. The goal is to support technologists in analyzing complex production data, identifying relevant process patterns, and deriving data-driven insights for quality optimization. This work defines a generalized workflow that enables technologists to move systematically from exploratory data analysis to hypothesis generation, experimental validation, and continuous process monitoring.

Various software platforms for data visualization and dashboard development exist, including Microsoft Power BI, Tableau, Qlik, and similar business intelligence solutions. The visual analytics environment presented in this work was implemented using Microsoft Power BI as the deployment platform. The contribution of this work does not lie in the development of a new visualization software, but in the design of a domain-specific visual analytics approach for aluminum casting process optimization, integrating data transformation, process–quality linkage, and interactive analysis workflows.

The increasing digitalization of industrial production processes has led to a continuously growing availability of process and production data. Numerous studies have addressed the analysis of such data as well as visual analytics approaches for supporting industrial decision-making. At the same time, a recent review [11] highlights that industrial applications are characterized by heterogeneous data sources and a strong need for domain knowledge. Although various visual analytics systems exist for industrial use cases, they often focus on individual aspects. One particular challenge involves combining data, expert knowledge, analytical methods, and interactive visualizations within reusable and integrated systems that support users in analyzing complex production data and making decisions.

In contrast to studies that address individual building blocks of Knowledge-Assisted Visual Analytics, the review [11] identifies a need for more comprehensive systems that combine data, expert knowledge, analytical methods, and interactive visualizations. Furthermore, the authors highlight the need for research on standardized and reusable architectures, improved integration of domain knowledge, and more comprehensive evaluation of developed systems under real industrial conditions.

To address these challenges, the present work makes the following contributions:

1. Development of an interactive visual analytics environment (“ADAM Process Data”) for the analysis of large volumes of industrial process data.
2. Integration of discretized process data, quality-related target variables, statistical analysis methods, and technological domain knowledge within a common analytical environment to support data-driven decision-making.
3. Definition of a generalized workflow that supports technologists from exploratory data analysis through hypothesis generation and validation in process trials to continuous process monitoring.
4. Long-term industrial validation of the proposed workflow, ranging from exploratory data analysis and process trials to the transfer of identified measures into industrial production and the evaluation of the resulting quality improvements.

2. Methods—From Time Series to Visual Insights: Interactive Analysis of Process Data

When datasets contain complex data with numerous parameters or are large in size, it becomes difficult to derive and extract meaningful information. This is particularly challenging in the present context, as the objective is not to examine individual features, but rather to consider entire time series that have been recorded during the production process [2,3].

2.1. Database

The visual analysis tool "ADAM Process Data", which will be described in more detail below, focuses on the visualization of casting process data. As previously mentioned, it is possible to cast up to six ingots in a single batch, with the casting process taking approximately two hours. During this period, a large amount of data is continuously recorded by hundreds of sensors and stored as time series. These sensors measure various parameters such as value, pressure, or speed in real time. The recorded data provides a detailed picture of the production process and forms the basis for optimization procedures [12]. The tool is used to make large, complex data volumes available in a clear and compressed form and to enable visual analyses to identify correlations with quality deviations that might otherwise remain hidden.

Due to confidentiality restrictions, exact dataset dimensions cannot be disclosed. The visual analytics environment is connected to several years of historical production and quality data from two industrial casting plants. Depending on the specific analysis objective, different subsets of the available data can be selected by the user, for example, specific alloys, product formats, production plants, or time periods. The database includes hundreds of continuously recorded process signals as well as quality information obtained from ultrasonic inspection of the final plates.

2.2. The Key Challenges the Tool Aims to Address

The processing of production data is associated with a number of challenges that must be overcome with the visual analysis tool. A central problem is the enormous amount of data collected over several years, including thousands of measured values per sensor and cast batch, which requires efficient storage and processing of the data. In addition, the process data is usually available as a continuous time series without any direct allocation to individual batches. This necessitates preliminary structuring of the data to identify relevant process sections. Depending on the specific question being addressed, it may also be necessary to consider different areas of the time series assigned to a particular batch. Some defects manifest only at the beginning of the casting process of a batch, in which case the initial part of the time series is relevant for analysis. For other defects, however, different time ranges may be more informative. This is discussed in more detail in the following example in Section 2.4.

Another challenge is the consideration of dynamic quality fluctuations that occur during the casting process, as the quality of an ingot or a batch can undergo changes during the production process. To illustrate this point, Section 2.4 examines a scenario in which an ingot at the beginning of casting, also referred to as the ingot foot, has a significant number of defects and is therefore classified as unsatisfactory. By contrast, at the end of the casting process, referred to as the ingot head, the ingot exhibits only a small number of defects and is rated as satisfactory.

Each batch is characterized by its own time series, which must be divided into meaningful sections to recognize specific correlations between process parameters and product quality. The subsequent sections will provide a more detailed description of this process.

Another crucial aspect is data transfer and pre-processing. The data originates directly from the production facilities and must be transferred to a cloud environment, where it is pre-processed prior to utilization for analyses. Ensuring the integrity of the data and its transformation into a standardized format is therefore of paramount importance. The successful handling of these challenges necessitates the deployment of specialized technologies and tools, which facilitate precise and scalable analysis.

2.3. Solutions to Overcome the Challenges

To address these challenges, a stable data connection to the cloud was established to facilitate continuous and automated processing of the large volumes of data, thereby ensuring users receive updated data on a daily basis. The data undergoes processing and allocation to respective batches on a daily basis, and a time-length transformation is also conducted. This involves assigning the time-related sensor data recorded during the casting of an ingot to a specific ingot length. The sensor position in the launder, the flow speed of the liquid aluminum, and the duration in special cleaning systems along the launder must be taken into account.

Another important step was the insertion of plausibility limits to identify and exclude incorrect or deviating data at an early stage. This procedure results in a significantly higher quality and precision of the analysis results.

The time series were then discretized [2,3,13–15]. In this process, the continuous time series were subdivided into manageable units (batch/ingot sections). This methodology enables the user to view production processes in a structured manner and analyze individual production steps. A detailed description of this process is provided in Section 2.5.

Finally, the process data was linked with quality data, in particular with a quality measure [3], which is described in more detail in Section 3.2, to determine the relationship between the production parameters and product quality. The use of this system enables well-founded decision-making, allowing for the implementation of specific measures to optimize production. A detailed explanation of this procedure can be found in Section 3.

2.4. Industrial Use Case: Non-Metallic Inclusions in Aluminum Casting

A particular technical issue in the foundry environment is selected to demonstrate the advantages of the visual analysis tool "ADAM Process Data" for the user. This example is used to illustrate the various visualization techniques in more detail. The discussion will focus on the advantages of discretized time series [13–15], particularly in the context of quality data.

The visual analysis tool "ADAM Process Data" is designed to be flexible, offering a range of applications to address a variety of potential research questions.

Currently, the tool's functionality encompasses data from two casting plants, with the potential for expansion to include additional casting plants.

The technical issue on the basis of which the visualization tool is described in more detail relates to the production of aerospace plates. During the production of aerospace material, process fluctuations during casting can lead to the occurrence of non-metallic inclusions. The subsequent production process, which extends over several weeks, involves a series of steps including homogenization, rolling, heat treatment and final quality control. During the quality check, a range of methods are employed, including ultrasonic testing, to detect non-metallic inclusions in the final product [4,5,7]. However, due to inherent limitations of the methods, defects can only be detected post-production. This could result in the rejection of the entire final plate if a significant number of non-metallic inclusions are detected.

By examining historical data, it is possible to identify factors that contribute to quality deviations during the casting process. This, in turn, enables countermeasures to be implemented to reduce the reject rate, to improve plant utilization and to shorten delivery times. The following sections provide a detailed explanation based on specific examples.

2.5. Discretization as a Transformation for Process Data Visualization and Analysis

As previously stated, two major challenges must be addressed. Firstly, consideration must be given to the dynamic quality changes that occur during the casting process, as

the quality of an ingot or a batch can change during the production process. Furthermore, there is a requirement for compact and clear visualization of continuous measurements recorded in high resolution during the casting process.

The duration of the casting process for a given batch is approximately two hours, during which hundreds of signals are recorded in high resolution. Analysis questions frequently pertain to the comparison of process data between multiple cast batches, with the objective of identifying differences. While it is possible to examine a single signal and compare it across two or three cast batches, this approach becomes impractical when attempting to compare several hundred cast batches.

To address these challenges, the process data of a batch was divided into batch sections, and different discretizations of each signal were calculated for each section (see [3,7,13–15]). Discretization of process data is the process of characterizing the signal by individual discrete values. This facilitates the identification of patterns and trends and the comparison of hundreds of cast batches. Discrete data can be presented in concise diagrams (for example, boxplots, violin plots) and more efficiently characterized by colors (for example, in scatter plots). This also makes it easier to identify outliers and trends for batches over several years.

The methodology applied follows the approach described in detail in [3] and is only briefly outlined here; for a comprehensive description, the reader is referred to [3], as this article focuses primarily on the application of the method within interactive visual analytics tools rather than on the methodological details themselves.

The schematic procedure for calculating discretization from a process signal is illustrated in Figure 3 [2,3,7]. The cast ingot is depicted schematically on the right-hand side, with the ingot foot (corresponding to the casting start) and the ingot head (corresponding to the casting end) having been sawn away and not processed further, as indicated by the dark grey area. The remaining ingot was then subjected to a rolling process, which was followed by a laboratory test of the resulting plates (light gray area in the cast ingot). Non-metallic inclusions, which are formed during the casting process, can be detected through the utilization of an ultrasonic test, and these ultrasonic indications are displayed as blue points. It is noteworthy that the majority of these indications do not constitute a critical issue, as they are very minor errors. However, indications that are too large in size have the potential to result in the rejection of the plate. As can be observed in general, the quality of this ingot may not be uniform throughout the casting process. In the schematic example in Figure 3, a greater number of indications were detected at the ingot's foot compared to its head.

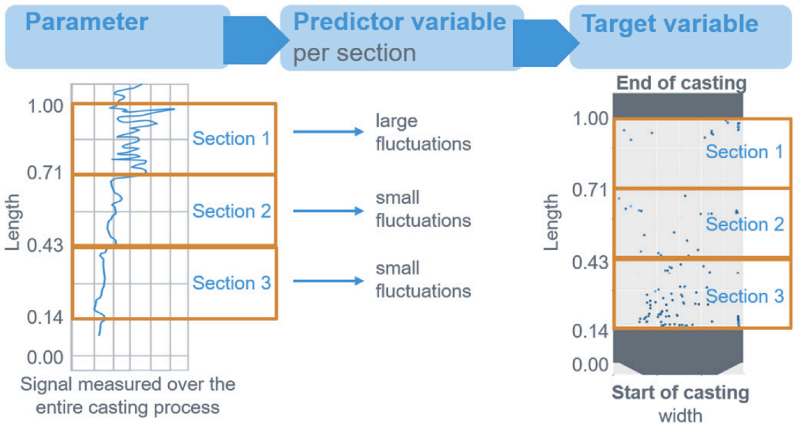


Figure 3. Schematic illustration of the discretization.

To address this variability, the ingot was segmented into different sections (i.e., Sections 1 to 3). In the following, the length of the ingot is normalized to the interval [0,1], a process which is also implemented for other dimensions. The initial section (S1) delineates the unstable part of the casting. The subsequent section (S2) encompasses the stationary phase, while the final section (S3) corresponds to the end of the casting. Nevertheless, alternative segmentations are also possible. The optimal segmentation is contingent on the specific question being addressed. In the example presented in Section 2.6.1, in addition to the sections S1, S2 and S3, the discretizations for sections S0 and S4 were also calculated. Section S0 corresponds to the initial casting stage (corresponding to the casting length interval which is milled away after casting), and section S4 delineates the final casting segment where the production of ingots is not uniform, depending on the specific product.

The left-hand side schematically illustrates a process signal after time-length transformation, allowing each signal value to be assigned to a specific position along the ingot length. Consistent with the segmentation of the ingot, the transformed process signal is partitioned into the same predefined sections. Subsequently, characteristic features such as mean values, fluctuation measures, and additional statistical descriptors are computed for each section (illustrated in the centre of Figure 3), yielding a discretized representation of the continuous process signal [13].

Further details on the discretization methodology can be found in [3].

This enables continuous process signals to be condensed into representative characteristics that can subsequently be compared across cast batches and evaluated with respect to quality deviations, such as non-metallic inclusions detected by ultrasonic testing. The example signal shown in Figure 3 is purely schematic and is only used to illustrate the discretization procedure. Any relationship between process signal characteristics and non-metallic inclusions is evaluated separately in the subsequent analyses.

2.6. Interactive Visual Analysis of Process Data: Techniques and Insights

The subsequent section of this paper will provide a detailed discussion of various visualization methods. Initially, the focus will be on visualizing process data over time, with a comparison of data from different plants and methods to identify outliers and trends. Subsequently, the visualization of disparate sections and their comparison will be presented. Finally, the paper will discuss the linkage of process data with production quality. In the discussion, the respective methods are presented using selected examples.

2.6.1. Process Data Visualization per Batch—Trend and Outlier Detection

The visualization over time, illustrated in Figure 4, enables the identification of trends and outliers. This facilitates the analysis of approximately 100 signals, with each signal recorded by two distinct casting plants. These plants are able to produce several hundred different alloys.

While the plants are essentially identical in principle, it is of particular interest to the user to compare analogue signals, such as values or rotor speeds, between the plants.

The following discussion refers to the example previously described. Four alloys (7xxx alloy family) intended for utilization in aerospace products were selected in the visualization. As an example, the mean value of a specific signal in the sections S1–S3, normalized to the interval [0,1] using min–max normalization, over several years, is presented. However, the user also has the option of selecting other sections using the 'Section' filter. In the following graphics, the time axis was anonymized, with the time stamps t1, t2, t3, etc., shown instead of specific time stamps.

The application incorporates filters that enable the user to adjust the time span to focus on the time ranges relevant to the analysis. The visualization furthermore offers additional filter options that allow the user to visualize different discretizations of the signal.

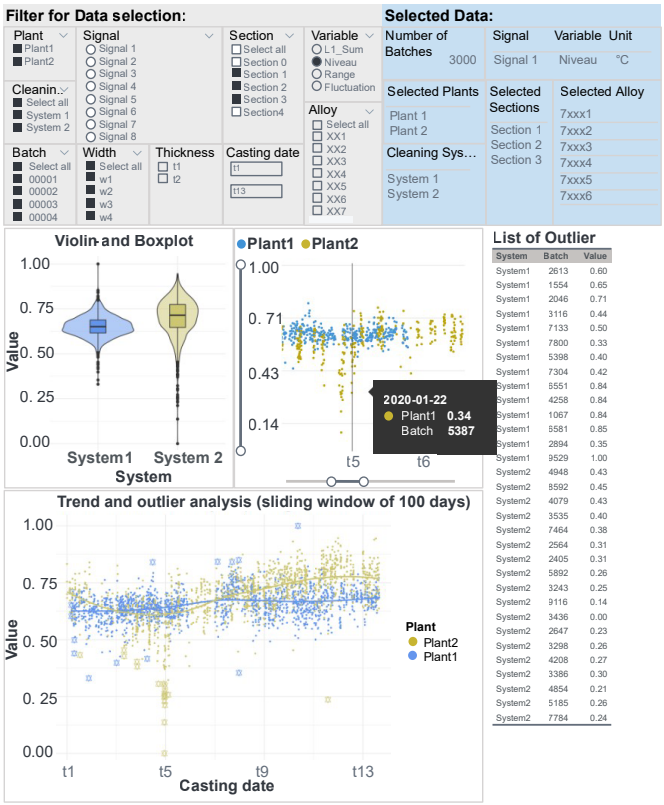


Figure 4. Process Data Over Time—Trend and Outlier (due to development trials) Detection.

This facilitates the comparison of equivalent values between the plants in the stationary casting area and provides a rapid overview of the value ranges in which the respective system operates, highlighting any differences between them. It offers process technologists a foundation for a better understanding of the processes and for making adjustments between plants to optimize product quality. Additionally, it facilitates a more efficient decision-making process when analyzing production data that lies outside the conventional process window.

As shown on the left of Figure 4, the visualization simultaneously presents a violin plot and a box plot for each system. This allows the user to rapidly identify the standard process window and the distribution of the signal, as well as to determine the presence of any differences between the plants. The visualization enables efficient communication of information regarding the process window, such as the median or the upper and lower quartiles, to the production department. The graph shows that plant 2 produces significantly higher values than plant 1 in the selected time range.

A visual examination of the graph entitled 'Trend and outlier analysis (sliding window of 100 days)' is sufficient to establish whether the observed behavior persists over the entire selected time range. The signal is represented by different colors for the two plants over time. The plotted line represents a trend curve of the data, which was determined using a locally weighted regression (standard for non-linear relationships) [13,16–19].

This visualization provides the user with direct insight into the development over the past few years. The analysis indicates a high degree of similarity in the values between the two plants until t7, as illustrated in Figure 4. Since t7, however, an increase in value can be observed at plant 2 that differs significantly from the original trends. In contrast, plant

1 had a relatively stable value. Finally, Section 3.3 provides a visual representation that focuses on the influence of the level of value in relation to product quality, particularly in the context of non-metallic inclusions.

In the graph entitled 'Trend and outlier analysis (sliding window of 100 days)' in Figure 4, it is possible to analyze outliers in addition to the trend. In this context, Tukey's method was applied, which identifies outliers using the interquartile range (IQR) and marks values outside the interval [Q1 − 3 × IQR, Q3 + 3 × IQR] as outliers [20]. The utilization of a sliding window facilitates the dynamic application of this calculation over temporal periods (e.g., 100 days), a method that is particularly well suited to process data exhibiting temporal patterns [13,16–19,21]. The presence of outliers is indicated in the graph by star markers. Furthermore, the user is provided with a list of outliers over time, which is presented in tabular form on the right-hand side and contains the associated batch number. This enables the user to rapidly identify batches that, while produced within the specified tolerances, fall outside the conventional process windows. Subsequently, these batches can be examined for any potential quality deviations.

To gain a deeper insight into the time series, the user may also make use of the upper right-hand graph. This offers a zoom function with sliders on the *x*-axis and *y*-axis. In the present study, a zoom was applied to the time period around t5, in which plant 2 showed some significant outliers. This provides the user with additional information about the exact date, the batch number and the exact value of the signal.

In this particular context, it has been demonstrated that signal discretization is beneficial. It allows the user to efficiently gather information regarding signals that were recorded during the casting process of one batch and to subsequently compare these with multiple others. In contrast, without signal discretization, the comparison of signals between multiple individual casts would not be possible. The analysis of individual time series is likely to be overwhelming for the user and hinder the identification of information and patterns. Discretization is therefore a significant advantage, as it makes it much easier to visually compare hundreds of cast batches.

In addition to the mean value, the user is also provided with other discretizations. These include, for example, a trend estimation of each signal for each casting process. Furthermore, the standard deviation is determined, as well as the range within which the signals fluctuate for each individual cast batch. This analysis is performed not only for the entire signal, but also for individual signal sections, as will be explained in more detail in the next section.

The visualization tool "ADAM Process Data" for the analysis of casting parameters incorporates interactive filters (illustrated by the grey marked area in the upper part of the graphic) that facilitate flexible and efficient data analysis. Additional intuitive filters, such as ingot format, alloy, or casting date, are straightforward to utilize, enabling the user to efficiently isolate relevant data. The integration of supplementary filters, including those pertaining to final plate quality (a subject to be addressed in detail subsequently), enables more in-depth investigations and specific analyses. The dynamic visualizations adapt immediately to changes made to the filters, which in turn promotes direct feedback.

Tooltips provide users with additional details on individual batches (e.g., equipment used, exact signal value and batch number) without cluttering the visualization, allowing for deeper investigation of individual batches. The additional export options improve usability by making data available to the user for presentations or reports.

The presence of clearly identifiable filter settings (illustrated by a blue marked area in the upper part of the graphic in Figure 4) further assists in keeping track of the applied filters, thus ensuring transparency in the data analysis.

2.6.2. Process Data Visualization per Batch Section

In the preceding section, the trend and outlier visualization for process data of an entire batch in the stationary casting area, with a comparison between two different plants, was demonstrated. The second visualization, which is also provided by the visual analysis tool 'ADAM Process Data' and which is presented in this section, is intended for data visualization and analysis from a single plant. In the example presented in Figure 5, the plant in question is plant 2.

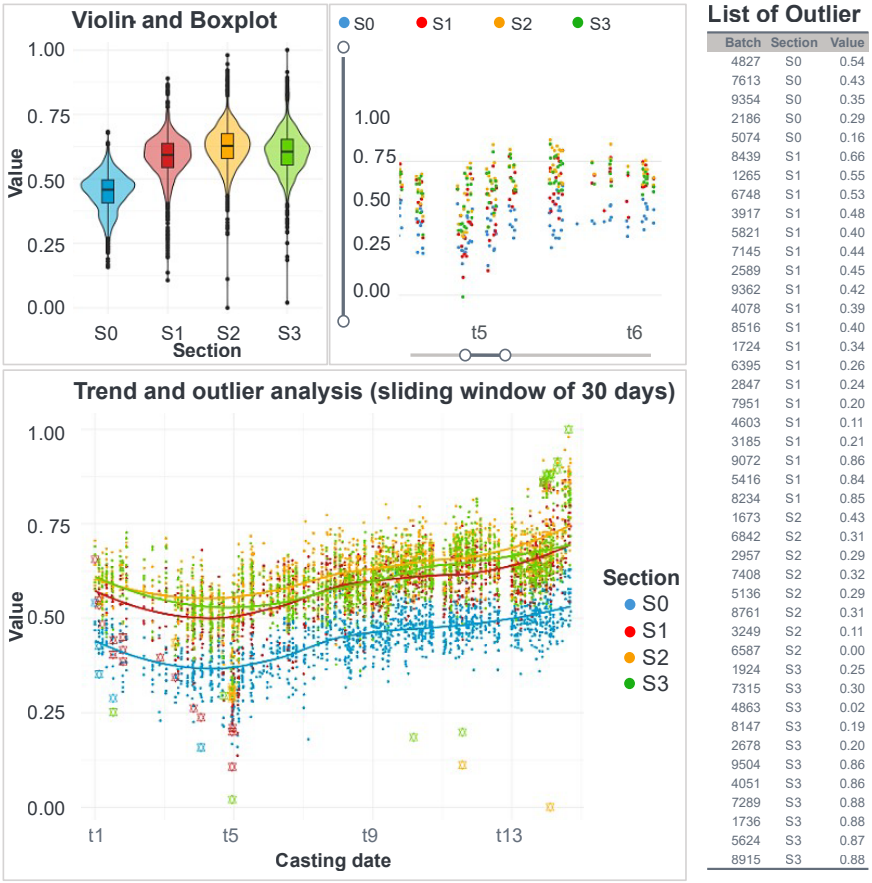


Figure 5. Process Data Over Time per batch section for plant 2—Trend and Outliers (due to development trials) Detection.

For the analysis of dynamic quality changes during casting (as the quality of an ingot or a batch can change during the production process), the signals of each batch were divided into several sections and discretized, as previously mentioned. The resolution of this discretization must be chosen according to the objective and the characteristics of the data to avoid loss of information.

In this example, a subdivision into four sections was chosen, with the first describing the start of the casting, the second the unstable part of the casting, the third the stationary phase and the fourth the end of the casting, as described in Section 2.5.

The violin/boxplot in the top left of Figure 5 compares the individual phases of the casting process (start, unstable phase, stationary phase, and end of casting) for the value signal from Figure 4 for plant 2. The difference between the phases is clearly visible. This is also evident in the graph 'Trend and outlier analysis (sliding window of 100 days)' [13,16–20]. The previously established trend towards higher values is evident in every section. Conclusions regarding the potential for optimization can be drawn based on

the connection with quality data, such as that pertaining to non-metallic inclusions, as will be described later in Section 3.3.

3. Results—Evaluating the Impact of Process Parameters on Product Quality Using Interactive Visual Analytics

In the preceding section, the potential of the developed visual analysis tool "ADAM Process Data" for the visual analysis of process data was presented. The section specifically addressed the benefits of discretizing process data for the visual analysis of multiple castings. In this section, the integration of process data with product quality data is illustrated using the example described in Section 2.4.

Section 3.1 first provides an explanation of how visualization techniques are already employed at AMAG to visualize product quality data. The subsequent sections then present a method that is implemented in the 'ADAM Process Data' visualization tool to analyze process data with product quality data. This method offers the user the opportunity to analyze influences on deviations in the quality of the final products. The following discussion is based on the example presented in Section 2.4.

3.1. Visualization of Product Quality Data

The prototype of the tool was presented in [9,10] as an interactive visualization tool for the data analysis of casting ingots. It consists of two main components: the visualization of the ultrasonic defects in the casting ingots with scatter plots and bar charts, as well as a filter function for various parameters (e.g., batch number, casting pit, plate thickness, indication size and alloy).

The prototype was further developed, and additional tools were implemented as shown in Figure 6. The visualization presented in Figure 6 has the capacity to illustrate the ultrasonic indications of a total of more than 10,000 ingots of several different alloys that have been produced at two different plants over a period of several years. In the filter selection shown in Figure 6, 25 ingots are selected as an example. Detected defects over the ingot length and width and also over the ingot thickness and width, are shown as a scatter plot. Furthermore, bar charts are displayed, which illustrate the mean number of defects per ingot segment. If the user selects a particular segment, it is highlighted in dark blue. In the example presented here, the last three segments along the length direction are selected, containing one critical defect shown in red. Colored points within the diagrams denote defect sizes. Red dots represent critical defects (class C), green dots represent non-critical defects (class A), and orange dots represent defects that can be critical if there are too many of them in one ingot and they are close to each other (class B). This color differentiation enables the user to rapidly assess the significance of the defects. Again, it should be noted that the illustration presents data from a limited number of batches, specifically only ten batches. For instance, if the user selects all data for an alloy, the specific distributions of the defects for this alloy are displayed. The scatter diagrams offer an intuitive representation of the distribution of non-metallic inclusions within the ingot, complemented by bar charts that provide additional information.

The interactivity facilitates the filtering and comparison of data, enabling the discovery of patterns. It is noteworthy that modifications made to one visualization are automatically synchronized with the others. Functions such as zoom, selection tools and highlighting facilitate exploratory analyses [22].

Another feature is the linking of forward and backward analyses: by selecting areas in the scatter diagrams or bar charts, users can examine specific patterns and recognize differences. A hover tool provides additional detailed information [23].

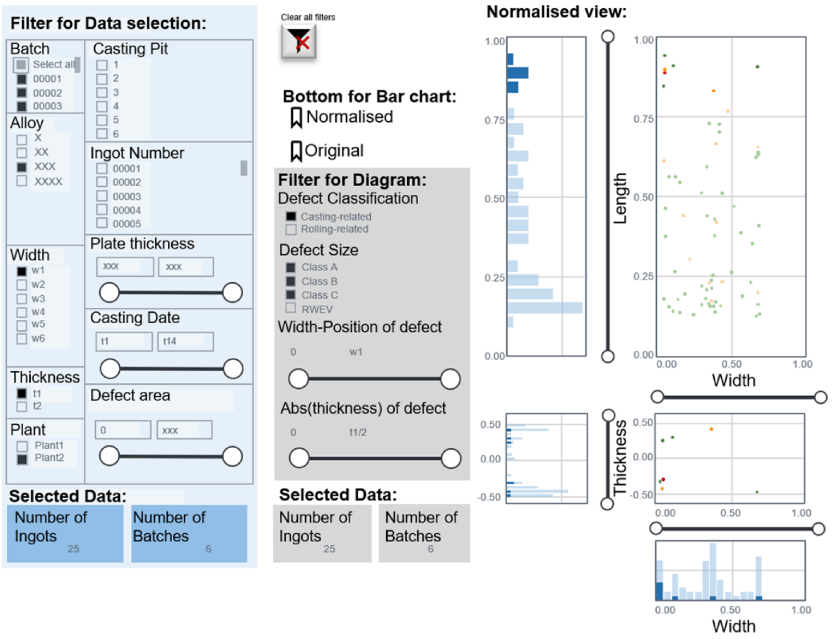


Figure 6. Visualization of Product Quality Data.

Two additional buttons have been added, allowing the user to switch between different views. The 'Normalized' view shows the normalized bar chart, meaning that the exact number of ingots produced is considered, so that ingots without defects are also taken into account when calculating the number of defects per section for the bar chart. The 'Original' view presents the bar chart exclusively in relation to ingots for which defects were detected. Depending on the research question, the user can switch between these views.

Furthermore, the user is provided with two distinct types of filters, distinguished by color-coding: blue and grey backgrounds. Filters with a blue background encompass filters such as alloy, plant, ingot width and thickness, which affect the selection of ingots displayed in the scatterplots and bar plots. The filters with a grey background do not affect the selected ingots, but only the indications shown in the graphics.

These include, for example, filters for size and position of indications, enabling the user to display indications of a specific size for all selected ingots. This makes it possible to analyze defect patterns of major defects, which are always relevant for product quality.

As demonstrated in Figure 6, the scatterplot provides a visual representation of the distribution of indications (of the 25 selected ingots) in relation to ingot thickness and thickness direction, relative to the position of the ingot. Additional bar charts are shown at the edge of the scatterplots, which show the number of indications in the respective area. This shows that the distribution of indications changes over the ingot length, and consequently also over the casting process.

At this point, it should be noted again that only 25 ingots have been selected in the figure, but the user has the option of viewing the distribution for all ingots, e.g., of an alloy or a plant, and analyzing the dependence of the distribution of the defects on the alloy or the plant.

It is also notable that the value is a parameter that changes during the casting process. This is illustrated in Figure 5, which presents the mean value of the signal in four distinct casting sections. In the initial casting phase (Section S0), the value is lower than in the stationary phase (Section S2).

The central question guiding this study is whether a correlation can be established between the value and the casting quality with regard to non-metallic inclusions. To this end, an analysis is performed in which the discretized signal values and the casting quality

are examined in relation to each other. However, it is first necessary to determine a measure (target variable) for each ingot section that describes the casting quality of an ingot after casting. This is outlined briefly in the following section. Further details can be found in [3,24].

3.2. Definition of a Target Variable for Quality Evaluation

In the context of conducting an ultrasound inspection, the precise locations and dimensions of non-metallic inclusions in each plate are determined. To identify the potential factors influencing the formation of such inclusions during the casting process, it is essential to develop a target variable that describes the quality of the ingot directly after casting, independent of any subsequent processing. Consequently, in a preliminary step, using the detected positions of indications in the plates, the original defect positions in the ingot are calculated. The primary challenge associated with this procedure pertains to the consideration of various sawing wastes and substantial dimensional changes that occur between process steps at different values and, consequently, at varying densities. It is imperative to recognize that these substantial dimensional changes can lead to significant errors in the recalculation of inclusion positions, even with minor deviations.

Once the calculation of the defect positions and volumes on the ingot has been completed, the calculation of the target variable can be initiated. A data scientist must overcome numerous challenges in this process. The methodology applied follows the approach described in detail in [3,24] and is therefore only briefly outlined here. In this context, a target variable is defined to quantify product quality based on non-metallic inclusions. Specifically, the target variable quantifies the sum of the indication areas [mm^2] per tested plate weight [t]. This will also be referred to as the quality measure defined by the following:

$$Q_i := \frac{\sum_j A_j}{W_i} \tag{1}$$

for an ingot i, where A_j is the indication area of a particular indication j in the ingot and W_i is the total weight of the tested plates produced from the ingot i.

Within a batch, there is a maximum of six ingots. However, it is possible for a single ingot in a batch to be of poor quality, even if the others are of very good quality. This phenomenon cannot be attributed to a batch-related variable. Rather, it is the result of an ingot-specific influencing variable. A Monte Carlo simulation is used [13,25] to identify such outlier ingots in a batch (see [24] for more details), because in the subsequent analysis, the focus will be exclusively on batch-related variables.

Finally, the target variable for the entire batch B (quality measure per batch) is defined as follows:

$$Q_B = median\{Q_i : i \in IB\} \tag{2}$$

where Q_i is the criterion of quality of an ingot i defined in Equation (1) and IB is the set of all ingots belonging to batch B, excluding outlier ingots.

The phenomenon of outliers can also be observed when considering ingot sections (referred to henceforth as 'outlier ingot (sections)'). The individual ingot/batch sections are defined here in three sections (S1, S2, S3). The first section (S1) describes the beginning of the casting, the second (S2) the stationary phase and the third the end of the casting. Section S0 is not relevant for the given question, as it is milled away directly after casting and is therefore not ultrasonically tested as described previously in Section 2.5.

The quality measure per ingot section and finally per batch section is thus defined in a manner analogous to the definition for the entire ingot or the entire batch (Equations (1) and (2)). However, instead of the entire ingot (or entire batch), only the respective length

section (S1, S2, S3) is considered (refer to [3,23,24] for a further detailed description of the methodology).

3.3. Evaluating Process–Quality Relationships Using Visual Analytics

In this section, a combined analysis of the preceding results is conducted to provide the user with an opportunity to evaluate the effect of specific characteristics of selected process data on the final product quality. The presentation of the results is once more based on the aforementioned example of non-metallic inclusions, but it is also possible to transfer this method to other characteristics of the final product quality.

While the methodology itself is presented in detail in [3,23,24], Figure 7 focuses on its implementation within the interactive visual analytics tool ADAM Process Data. In contrast to the methodological description in previous work, the figure demonstrates how the approach is applied in practice, enabling users to interactively analyze and interpret the relationship between process parameters and product quality.

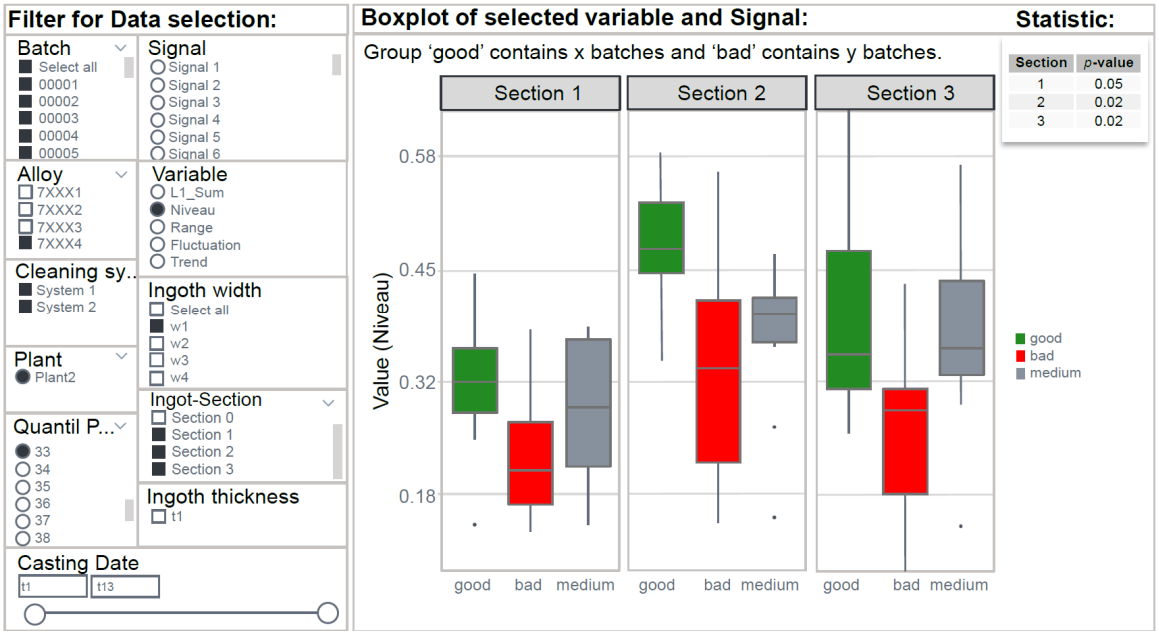


Figure 7. Visualization of discretized process data in relation to product quality data (data until May 2024).

The visualization in Figure 7 is divided into three components. The first component, located on the left, presents the filter options available to the user. Next to these options is a boxplot, which provides a visual representation of the results. The third component, situated on the far right, performs an assessment of the results using statistical methods. In the following, a more detailed description of each component is provided.

In Figure 7, all batches of a specific alloy with the same format for aerospace plates for several years were selected. The discretized level value is shown normalized in the middle graphic of Figure 7, with the three parts of the graph describing the results for the three batch sections. Section 0 (S0) was not analyzed in this case because this section is cut away immediately after casting in aerospace products for quality reasons and is therefore not ultrasonically tested, as mentioned before. However, for other products or issues, this area can be relevant for product quality.

Depending on the research question, other parts of the process signals can be relevant for quality. For this reason, the user also has a filter option on the left-hand side to select the sections relevant to the problem under discussion.

The quality measure per batch section (Equation (2)) has been categorized into three groups in each of the three sections of the graph in Figure 7. The three groups describe the quality of the final batch with regard to the previously defined quality measure per batch.

The category 'good' includes the q = 33% of the selected batches with the lowest quality measure per batch section. These are the batches with the best quality in terms of non-metallic inclusions (that means the fewest) in the respective section. The parameter q denotes the quantile parameter used for defining the category boundaries. The 'bad' category is defined by the 33% of batches with the highest quality measure per batch section, while the 'medium' category encompasses the remaining batches per section. It should be noted here that the same batches do not always have to be in the 'good', 'medium' and 'bad' categories in each section, as the quality of batches can change significantly over the casting length, as previously mentioned. The user can select the parameter q themselves by utilizing the filters located on the left-hand side ('Quantile Parameter'). In addition, the heading of the graph in the center shows the user how many batches are in the 'good' and 'bad' categories to give users an impression of the significance of the boxplots. It is important for users to know that with a limited sample size, the significance of the boxplots is reduced.

As illustrated in Figure 7, the samples classified as 'good' have consistently higher level values compared to those categorized as 'bad'. To provide the user with a more comprehensive assessment of the results, a statistical analysis is also displayed on the right-hand side under the 'Statistics' tab. Therein, a comparison is made between the two groups, 'good' and 'bad', for each section. A non-parametric test was implemented (R package npmv introduced in [26]), which provides the user with the p-value for the comparison between the two groups 'good' and 'bad'. Differences between the quality groups are assessed using the non-parametric comparison test nonpartest implemented in the R package npmv [26]. The method compares the distributions of the investigated groups without assuming multivariate normality and is suitable for small sample sizes as well as mixed variable types. This is particularly useful in this context, because the user naturally cannot perform a complete analysis of the selected data based on the requirements of the implemented test but still requires an indication of the significance of the boxplots. The p-value, or significance value, is used in statistics as a measure describing the probability of obtaining a more "extreme" value in the direction of the alternative hypothesis, assuming that the null hypothesis actually holds. A "small" p-value (less than a certain significance level α; generally, $\alpha \coloneqq 0.05$) leads to rejection of the null hypothesis [27].

This selected example demonstrates significant discrepancies in the level values of the signal between batch sections classified as 'good' and 'bad'. The user now has the option of examining the entire data set for additional time periods, thereby eliminating the possibility of a temporal influence. Furthermore, the investigation can be extended to other alloys and formats. The outcome of this investigation is that, independent of the specific time frame, there is a substantial difference in the level values of the 'good' and 'bad' categories, particularly for the given alloy illustrated in Figure 7. In addition, consistently higher values can be observed in the category 'good'.

This observation led to the decision to undertake a more comprehensive investigation, following the initial visual analysis of the distribution displayed in Figure 6, which already indicated the potential impact of the value on final quality. It was defined as a general procedure that, if signal behavior is observed that could potentially affect product quality, closer investigations should be conducted prior to implementing any measures to modify existing production processes.

It should be noted that the presented statistical analyses are primarily intended to support hypothesis generation and process understanding within the visual analytics

workflow. The same dataset was used for visual exploration and subsequent statistical comparison of quality groups. Consequently, the reported *p*-values should be interpreted with appropriate caution, particularly with respect to potential multiple-testing effects and the risk of over-interpreting exploratory findings. For this reason, statistical significance was not considered sufficient evidence on its own. Instead, the identified relationships were subsequently assessed by domain experts and validated through dedicated process trials and industrial implementation.

3.4. Further Visual Investigation

To investigate the behavior described in more detail, the indication distribution was examined more closely once again. In Figure 8, the defect distribution over the ingot thickness is shown. The same batches were selected for this representation as in Figure 7 for batch Section S2. The normalized level values in Section 2 were divided into four groups.

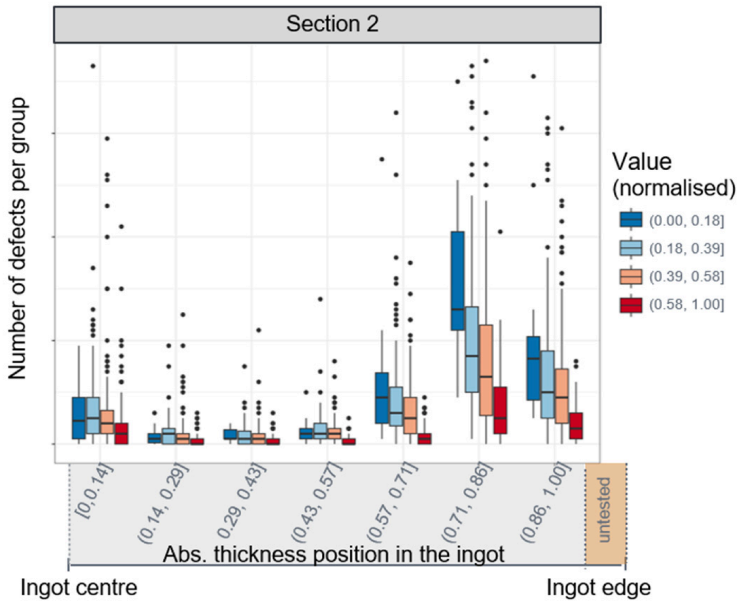


Figure 8. Defect distribution per value group.

Due to the necessity for technical symmetry, the thickness position of defects in Figure 8 was flipped at the center of the ingot and subsequently divided into groups. It should be noted at this point that the thickness coordinates are also normalized. This implies that for an ingot thickness of x mm (with the exception of the untested area marked orange in the graphic), the center in the thickness direction is 0 and the top and bottom edges are 0.5 and −0.5, respectively. The group '[0, 0.14]' contains, for example, all defects with thickness position coordinates at most $0.14 \times x/2$ mm from the center of the ingot.

The group '[0.86, 1.00]' in Figure 8 contains all defects with thickness position coordinates that are more than $0.86 \times x/2$ mm and at most $x/2$ mm away from the center of the ingot. These are located at the very edge of the ingot. The number of indications per ingot was now calculated for the individual value and thickness position groups and displayed in a box plot (Figure 8). It can be observed that the number of defects per ingot is lowest at each thickness position when the value is highest. In addition, there is a clear trend towards fewer indications with increasing value in each thickness group.

A technical explanation for this behavior is that at higher values, the flow behavior of the liquid aluminum when it is poured into the ingot is different from that at lower values. As a result, the non-metallic inclusions are positioned more at the edge of the ingot

at higher values. For reasons related to product quality, the edge of the ingot is milled away for aerospace products; thus, non-metallic inclusions in these areas are also milled away.

3.5. Impact—Process Change and Monitoring Using Visual Analytics

As all the evaluations indicated a positive influence of higher values, it was decided to run trials with specifically higher values outside the usual process window [8]. The visual analysis tool "ADAM Process Data" can be used for trial monitoring and continuous evaluation by the technologists, as all relevant data, from the measured signals to information on the final product quality, is available to the user in this visual analysis tool on a daily basis. This enables the technologists to assess the effects of the test batches and the final product quality immediately.

Figure 9 again shows the visual analysis tool from Figure 7, but now with data up to October 2024, i.e., including the area where the value was deliberately raised on a trial basis. It is clearly visible that the batches in the 'good' category had a much higher level value than the batches in the 'bad' category. The presentation of the results of the non-parametric test in the 'Statistics' tab also shows that the *p*-values are close to 0 in all sections. This shows very clearly that the controlled higher value results in an increase in final product quality in terms of non-metallic inclusions.

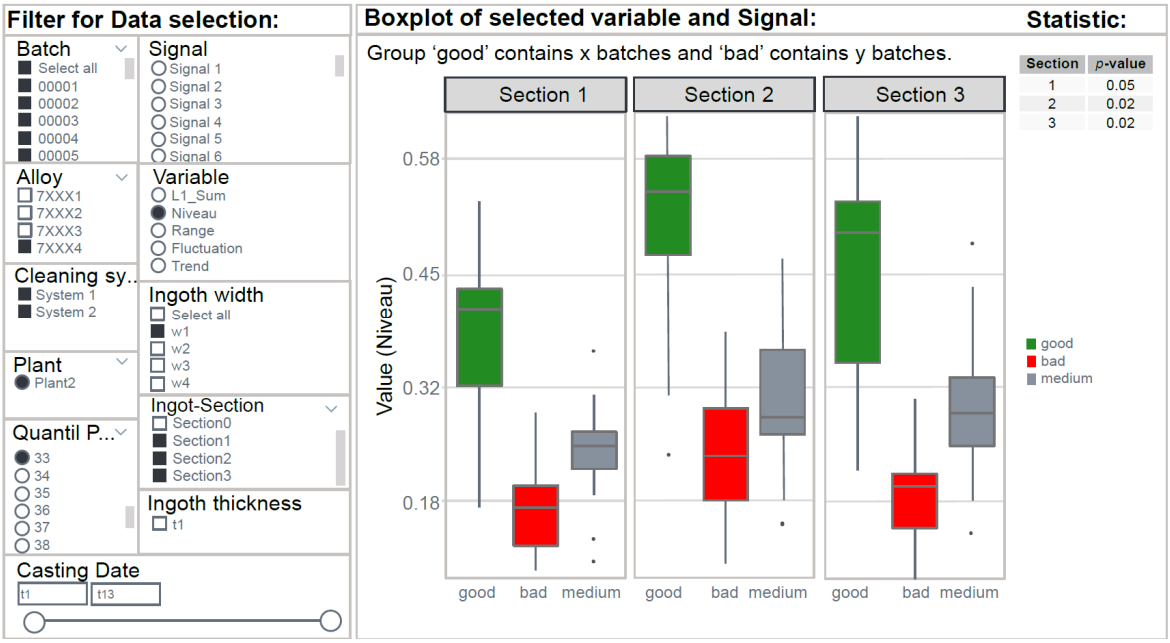


Figure 9. Visualization of discretized process data in relation to product quality data (data until October 2024).

This experimental process change was then incorporated into the production process. This resulted in a reduction in the rejection rate (regarding non-metallic inclusions) of over 70% for the alloy considered in the previous analyses. The trial changeover started in t13. This resulted in an immediate 70% reduction in the scrap rate until t14. This shows that visual and statistical analyses of large amounts of data can make a significant contribution to improving quality in industrial processes [8].

The resulting workflow, enabled by the visual analytics tool, can be summarized as follows:

First, technologists use the interactive visual analysis tool to explore process data and identify relevant patterns. Based on this exploration, domain experts formulate hypotheses regarding potential relationships between process parameters and product quality. Together

with data science expertise, the hypotheses are analyzed and subsequently tested through targeted process trials in production. During these trials, the visual analytics tool enables continuous monitoring and evaluation of both process data and resulting product quality. This allows technologists to directly assess the effects of parameter adjustments in an iterative and data-driven manner.

If the results confirm the hypothesis, the corresponding process changes can be implemented in production, leading to measurable improvements in product quality. Following implementation, the tool continues to support long-term monitoring by facilitating trend analysis and outlier detection, ensuring the stability and sustainability of the process improvements.

3.6. Adapting the Approach to Different Industrial Processes

However, application-specific adaptations are required when transferring the approach to different processes. In particular, the definition of the target variable, the segmentation of the process, and the discretization strategy must be aligned with the underlying process characteristics and analysis objectives. General considerations regarding the definition of target variables in multi-stage industrial production processes are discussed in [28], while further examples of application-specific data transformation approaches are presented in [29].

The following section describes the general approach one can take when developing a segmentation and discretization strategy for a specific application. This is illustrated for the casting and rolling process in Figure 10.

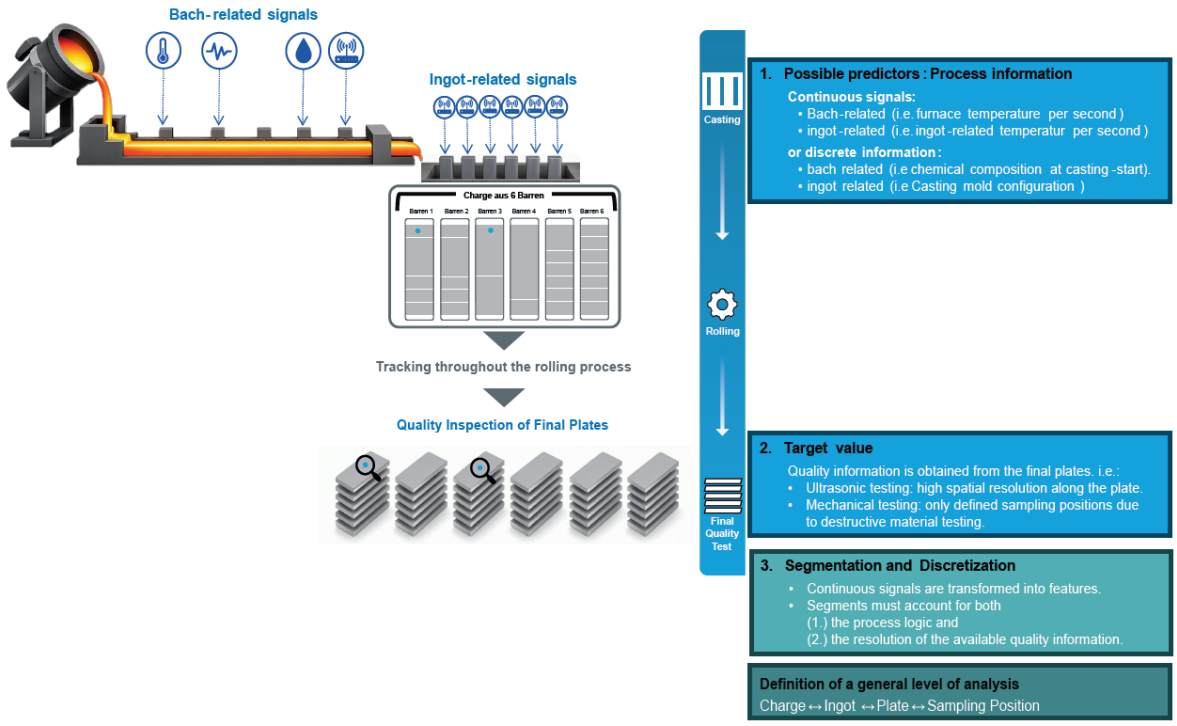


Figure 10. Availability and Resolution of Process and Quality Information.

1. Availability and Resolution of Process and Quality Information

In industrial production processes, both process information and quality information may be available in different forms and at different levels of aggregation. Therefore, before analyzing relationships between process variables and product quality, it is essential to understand which information is available and at which level it is represented.

- Continuous Process Signals

Continuous process signals are recorded throughout the entire production process (the left side of Figure 10). A distinction can be made between batch-related signals and ingot-related signals. Batch-related signals describe the production of an entire batch, which may consist of several ingots. A typical example is the furnace temperature, which is continuously measured during production and applies to the entire batch. In contrast, ingot-specific signals can be directly assigned to an individual ingot.

- Discrete Process Information

In addition to continuous signals, many process variables are already available in discrete form. These include categories or system configurations that do not require further discretization.

An example of discrete batch-related information is the chemical composition of the melt. During the production of a batch, melt samples are typically taken at predefined time points, for example, at the beginning, middle, and end of the casting. The resulting analytical values describe the entire batch from which several ingots are cast.

Ingot-specific discrete information, on the other hand, refers directly to a single ingot. Examples include the applied mold configuration or specific ingot-related equipment.

2. Target value

The target variable describes the final product quality and is based on quality measurements performed on the final plates. Consequently, quality information becomes available only after completion of the entire production process, i.e., after casting, subsequent rolling, and final quality inspection (see Figure 10).

Depending on the inspection method, the spatial resolution of the quality information may vary significantly.

In the case of ultrasonic testing, the finished plates are inspected nearly over their entire length. Detected indications can subsequently be traced back to their original positions within the ingot. As a result, quality information is available with a high spatial resolution along almost the entire ingot length. This enables a comparatively flexible segmentation strategy along the ingot.

The situation is different for mechanical testing. Mechanical properties such as tensile strength, yield strength, or elongation are determined by destructive material testing at predefined sampling positions. Consequently, the quality information is only available at a limited number of discrete sampling locations and not continuously along the entire product length. In such cases, the segmentation strategy must also account for the sampling concept.

3. Segmentation and Discretization

Discretization is only required for continuous process signals. Its objective is to transform continuous signal profiles into a finite set of representative features, such as mean values, standard deviations, trends, minima, or maxima. This allows complex time series to be converted into comparable and interpretable variables that can subsequently be visualized, linked to quality information, and utilized in statistical analyses, feature-importance investigations, machine-learning models, or predictive analytics.

Segmentation refers to the decomposition of a process into quality-relevant sections. It is guided both by the process behavior and by the availability and resolution of quality information. In the present application, for example, segmentation is performed along technologically relevant, distinct regions of the casting process, such as the casting start, unstable phase, stationary region, and casting end.

The challenge of segmentation lies in linking process information with target variables. Process information and target variables may be available at different aggregation levels, for example, at the level of an entire batch, a single ingot, a plate, or even individual sampling positions within a plate.

In addition, process and quality information often differ with respect to their spatial and temporal resolution. Certain process variables may only be available for an entire batch, whereas product quality may be determined by individual ingots, individual plates, or only at selected sampling positions. Conversely, continuous process signals may be available at a high temporal resolution, while the target variable may only be available as a single quality value per batch, ingot, or product.

The segmentation strategy must account for these different levels of information. On the one hand, segments should represent technologically meaningful process regions. On the other hand, they must be defined in a way that enables consistent linkage with the available quality information. Consequently, the choice of segmentation is determined not only by process behavior but also by the definition, availability, and resolution of the target variable.

The overall objective is to transform process data and quality information into a common level of analysis so that reliable relationships between potential influencing factors and final product quality can be identified and interpreted.

4. Discussion—A Generalized Workflow for Visual Analytics in Process Optimization

Collaboration between different specialist areas is a crucial aspect in the development of such tools, as data only becomes meaningful in the context of specific application areas. The integration of expertise from different disciplines is therefore of great importance, as it helps to identify relevant issues, interpret analyses correctly, and make well-founded decisions in practice. This also requires the combination of statistical methods, technical expertise and domain-specific know-how to effectively optimize processes and develop innovative solutions. The exchange of different perspectives is therefore crucial to ensure the effectiveness of data-based findings.

In Figure 11, the workflow and interaction between data science, visual analytics and technology is illustrated. Technological knowledge was a central component in the development and application of the visual analysis tool ‘ADAM Process Data’. This tool assists technologists in identifying new hypotheses for optimizing existing processes. The validity of these correlations is verified using well-established mathematical and statistical methods prior to the implementation of optimization measures. The validation process is an essential component of scientific research, as it ensures the accuracy of the underlying analyses and prevents the drawing of incorrect conclusions. In the context of industrial applications, the delivery of robust statistical results is of particular importance, as this helps avoid unnecessary adjustments and incorrect decisions [24].

Due to the specialized nature of the application, the visual analytics environment was developed in close collaboration with its intended users, namely product technologists responsible for process analysis and optimization. Regular meetings were conducted throughout the development process to discuss requirements, evaluate intermediate implementations, and prioritize additional functionality. The analysis tasks addressed require substantial domain-specific process knowledge and are therefore performed by only a small group of experts within the organization. Following deployment, structured expert interviews were conducted with the active users of the system. A questionnaire was used as a discussion guide to collect feedback on practical benefits, usability aspects, and future development opportunities.

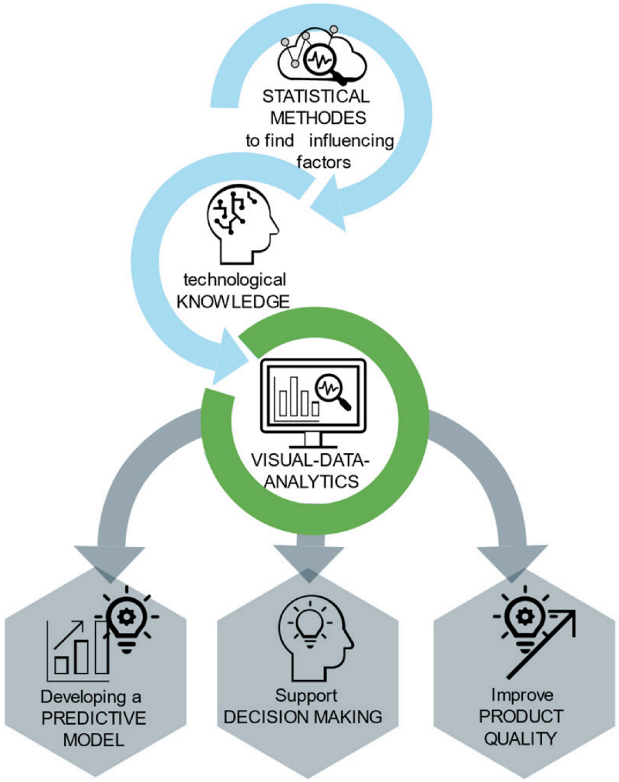


Figure 11. Workflow and interaction between different scientific fields.

The interviewees particularly highlighted the ability to explore large process datasets and to identify and discuss potential process–quality relationships. They further revealed strong interest in extending the system to additional casting devices and integrating further process signals and visualization capabilities. In addition, the introduction of simplified entry-level views was suggested to facilitate onboarding of future users with different levels of experience. The resulting feedback was incorporated into the development roadmap and forms the basis for several planned extensions of the system.

Beyond the qualitative user feedback, the practical relevance of the proposed visual analytics workflow is demonstrated by its successful application in industrial process optimization. Insights derived from the visual analytics process contributed to process changes that reduced the rejection rate (regarding non-metallic inclusions) by more than 70%.

5. Conclusions and Further Work

This work addresses two major challenges that frequently arise in industrial production processes dealing with large volumes of high-resolution data. First, the analysis of quality fluctuations during the production process is considered. Second, the visualization of continuous high-resolution measurements recorded throughout the process is addressed.

A central contribution of this work is the development of an interactive visual analytics approach that translates complex process time series and product-quality-related characteristics into a form that can be directly explored and interpreted by technologists in practice. While the underlying methodology for discretization and target variable definition has been described in earlier work (see, for example, [3,23,24]), the primary focus of this paper lies on its practical application within an interactive visual analytics environment.

A key contribution is the combination of discretized process data, a process-related quality measure, and interactive visualization within the visual analytics tool. This integrated approach enables users to systematically explore large numbers of production

batches, identify trends and outliers, and analyze process–quality relationships in an intuitive and application-oriented way. In this sense, the contribution of this work lies in transforming methodological concepts into an operational tool for industrial decision support and in defining a generalized workflow that enables technologists to move systematically from exploratory data analysis to hypothesis generation, experimental validation, and continuous process monitoring.

The industrial use case presented in this work illustrates the practical impact of the proposed approach. The interactive visual analysis tool revealed that certain process parameters have a particularly strong influence on product quality. Building on these insights, future research will address the development of real-time prediction models to enable the targeted and dynamic control of these parameters during production [29].

In addition to these model-based extensions, future work will focus on improving the calculation of the target variable [29] and further enhancing the capabilities of the visual analytics tool. In particular, advanced visualization techniques, such as the simultaneous analysis of multiple process signals (e.g., using parallel coordinates), will be investigated in order to better understand combined effects on product quality. In addition, improved interaction concepts will be developed to allow users to explore the full temporal evolution of process signals more efficiently.

The current version of the visual analytics environment is deployed for two similar casting devices and is therefore used by a relatively small group of domain experts. Based on the positive practical experience and user feedback obtained during the development process and the expert interviews, an extension of the approach to additional casting devices is currently planned. This will substantially increase the potential user base and enable broader user-centered evaluations in future work.

Furthermore, future developments will focus on simplified entry-level views that facilitate onboarding of new users while maintaining advanced analytical functionality for experienced users.

Beyond the presented use case, the approach is transferable to other industrial domains where complex quasi-continuous process data must be linked to product quality and where interactive exploration is essential for expert decision-making.

Overall, this work demonstrates that the combination of data transformation methods, interactive visualization, and domain expertise provides a powerful and transferable approach for analyzing complex production processes and enabling effective, data-driven quality optimization in industrial practice.

Author Contributions: Conceptualization, M.L.S.; methodology, M.L.S., A.G., S.N. and W.T.; software, M.L.S.; validation, M.L.S., A.G., S.N. and W.T.; formal analysis, M.L.S.; investigation, M.L.S., A.G., S.N. and W.T.; resources, M.L.S.; data curation, M.L.S.; writing—original draft preparation, M.L.S.; writing—review and editing, M.L.S., A.G., S.N. and W.T.; visualization, M.L.S.; supervision, M.L.S.; project administration, M.L.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets presented in this article are not readily available because the data are part of an ongoing study and to allow for the commercialization of research findings.

Acknowledgments: Wolfgang Trutschnig gratefully acknowledges the support of the WISS 2025 project 'IDA-lab Salzburg' (20204-WISS/225/197-2019 and 20102-F1901166-KZP). During the preparation of this manuscript/study, the authors used Microsoft 365 Copilot (M365 Copilot) for the purposes of language translation/editing/refinement and the creation of graphics for Figure 10. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors have reviewed the journal's guidelines regarding competing interests and potential commercial interests. Three of the authors are employees of Austria Metall GmbH and AMAG Casting GmbH. The study describes a methodology that is used internally within the company but is neither commercialized nor associated with any patents, licenses, financial interests, or other commercial benefits related to the publication. The authors confirm that there are no additional competing interests or commercial interests that require declaration.

Abbreviations

The following abbreviations are used in this manuscript:

IQR	Interquartile Range
Q1, Q3	Quartile (Tukey Method)
p-value	Significance Value
α (alpha)	Significance Level
S0, . . ., S4	Ingot/Batch Section
ADAM	Aluminum production Data Analysis and Monitoring for Process Data

References

1. Lee, J.; Kao, H.-A.; Yang, S. Service innovation and smart analytics for industry 4.0 and big data environment. *Procedia CIRP* **2014**, *16*, 3–8. [CrossRef]
2. Schreyer, M.; Gerber, A.; Neubert, S. Data Analysis of Production Data for continuous casting of Aluminum rolling ingots. *Key Eng. Mater.* **2023**, *968*, 95–101. [CrossRef]
3. Tschimpke, M.J.; Trutschnig, W.; Schreyer, M.; Neubert, S.; Gerber, A. Statistical methods for prediction of the quality of aluminium ingots. *Key Eng. Mater.* **2023**, *968*, 109–116. [CrossRef]
4. Vasudevan, A.K.; Doherty, R.D. *Aluminum Alloys–Contemporary Research and Applications*; Elsevier: Amsterdam, The Netherlands, 2012.
5. Prillhofer, B.; Antrekowitsch, H.; Böttcher, H.; Enright, P. Nonmetallic inclusions in the secondary aluminium industry for the production of aerospace alloys. In *Proceedings of the Light Metals 2008: Proceedings of the Technical Sessions Presented by the TMS Aluminum Committee at the TMS 2008 Annual Meeting and Exhibition, New Orleans, LA, USA, 9–13 March 2008*; TMS: Pittsburgh, PA, USA, 2008; p. 603.
6. Suschnigg, J.; Ziessler, F.; Brillinger, M.; Vukovic, M.; Mangler, J.; Schreck, T.; Thalmann, S. Industrial production process improvement by a process engine visual analytics dashboard. In Proceedings of the 53rd Hawaii International Conference on System Sciences, Maui, HI, USA, 7–10 January 2020; pp. 1320–1329.
7. Schreyer, M.; Tschimpke, M.; Neubert, S.; Gerber, A.; Trutschnig, W. Statistical methods for measuring and predicting the quality of aluminum slabs. *Light Metal Age* **2023**, *81*, 46–53.
8. Neubert, S.; Schreyer, M.; Gerber, A.; Pucher, P. Umsetzung von Big-Data-Analysen zur Verbesserung der Walzbarrenqualität. *ALUMINIUM* **2023**, *10*. Available online: https://epaper.metal.one/de/profiles/d416de48eeac-epaper-metal-one-aluminium-journal-blech-nc-fertigung/editions/int-aluminium-journal-10-2023/pages/page/18 (accessed on 15 September 2026).
9. Jekic, N.; Mutlu, B.; Faschang, M.; Neubert, S.; Thalmann, S.; Schreck, T. Visual analysis of aluminum production data with tightly linked views. In *EuroVis*; Eurographics: Montreal, QC, Canada, 2019; pp. 49–51.
10. Jekic, N.; Schreyer, M.; Neubert, S.; Mutlu, B.; Schreck, T. Visual Data Analysis of Production Quality Data for Aluminum Casting. In Proceedings of the 54th Hawaii International Conference on System Sciences, HICSS 2021, Kauai, HI, USA, 4–9 January 2021; pp. 1530–1605.
11. Böck, A.J.; Größbacher, S.; Vrablicz, J.; Stoiber, C.; Rind, A.; Suschnigg, J.; Schreck, T.; Aigner, W.; Wagner, M. KAVAI: A Systematic Review of the Building Blocks for Knowledge-Assisted Visual Analytics in Industrial Manufacturing. *Appl. Sci.* **2025**, *15*, 10172. [CrossRef]
12. Steven, M.; Klünder, T.; Lindstädt-Dreusicke, N. *Big Data: Anwendung und Nutzungspotenziale in der Produktion*, 1st ed.; W. Kohlhammer GmbH: Stuttgart, Germany, 2020.

13. R Core Team. *R: A Language and Environment for Statistical Computing*; The R Foundation for Statistical Computing: Vienna, Austria, 2022.
14. Walz, G. *Lexikon der Mathematik*, 1st ed.; Spektrum Akademischer Verlag: Mannheim/Heidelberg, Germany, 2000.
15. Barnert, S.; Delbrück, M.; Eis, R.; Fischer, N.; Greulich, W.; Heinisch, C.; Nagel, S.; Radons, G.; Schilling-Benz, L.; Schüller, J. *Lexikon der Physik*; Spektrum Akademischer Verlag: Mannheim/Heidelberg, Germany, 1998.
16. Wickham, H.; Navarro, D.; Pedersen, T.L. *ggplot2: Elegant Graphics for Data Analysis*, version 3.4.0; Springer: Berlin/Heidelberg, Germany, 2023.
17. Wickham, H.; François, R.; Henry, L.; Müller, K.; Vaughan, D. *dplyr: A Grammar of Data Manipulation*, version 1.1.3; CRAN: Vienna, Austria, 2023.
18. Microsoft Corporation. *Microsoft Power BI*, (Version: 2.158.1177.0 64-bit); Microsoft: Redmond, WA, USA, 2025.
19. Hester, D.; Wickham, H. *slider: Sliding Window and Rolling Calculation Functions*, version 0.3.1; CRAN: Vienna, Austria, 2023.
20. Tukey, J.W. *Exploratory Data Analysis*; Addison-Wesley: Boston, MA, USA, 1977.
21. Cormen, T.H.; Leiserson, C.E.; Rivest, R.L.; Stein, C. *Introduction to Algorithms*, 3rd ed.; The MIT Press: Cambridge, MA, USA, 2009.
22. Cleveland, W.S. *The Elements of Graphing Data*; Hobart Press: Murray Hill, NJ, USA, 2009.
23. Heer, J.; Bostock, M.; Ogievetsky, V. A Tour through the Visualization Zoo. *Commun. ACM* **2010**, *53*, 59–67. [CrossRef]
24. Schreyer, M.; Tschimpke, M.; Gerber, A.; Neubert, S.; Trutschnig, W. Applied Statistics in Industry: Defining an Appropriate Target Variable and Analysing Factors Affecting Aluminium Ingot Quality. In *Combining, Modelling and Analyzing Imprecision, Randomness and Dependence. SMPS 2024*; Ansari, J., Fuchs, S., Trutschnig, W., Lubiano, M.A., Gil, M.A., Grzegorzewski, P., Hryniewicz, O., Eds.; Advances in Intelligent Systems and Computing; Springer: Cham, Switzerland, 2024; Volume 1458.
25. Kroese, D.P.; Brereton, T.; Taimre, T.; Botev, Z.I. Why the Monte Carlo method is so important today. *WIREs Comput. Stat.* **2014**, *6*, 386–392. [CrossRef]
26. Burchett, W.W.; Ellis, A.R.; Harrar, S.W.; Bathke, A.C. Nonparametric Inference for Multivariate Data: The R Package npmv. *J. Stat. Softw.* **2017**, *76*, 1–18. [CrossRef]
27. Sachs, L.; Hedderich, J. *Angewandte Statistik: Methodensammlung mit R*, 8th ed.; Springer Spektrum: Berlin/Heidelberg, Germany, 2018.
28. Schreyer, M. Key Concepts and Challenges in Industrial Processes with a Focus on Defining Target Variables in Multi-Stage Production and Batch Data. In *Data Mining—Foundational Concepts and Cutting-Edge Advancements*; Sultana, T., Ed.; IntechOpen: London, UK, 2025.
29. Schreyer, M.; Gerber, A.; Neubert, S.; Simon, P. Enhancing Product Quality in High-Variant Manufacturing: Combining Physics-Based Simulations and Data Science for Target Variable Estimation in an IoT- and Machine Learning-Driven Context. *Sensors* **2026**, *26*, 830. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.