

processes



Article

Optimization of Manufacturing Processes Using AI-Based Advisory Systems: Casting Application

Sofija Milicic, Amir M. Horr, Stefanie Elgeti, Manuel Hofbauer and Rodrigo Gómez Vázquez

Special Issue

Artificial Intelligence in Process Innovation and Optimization

Edited by

Dr. Leonardo Rosa Ribeiro da Silva and Dr. Luiz Eduardo dos Santos Paes



<https://doi.org/10.3390/pr14162623>

Article

Optimization of Manufacturing Processes Using AI-Based Advisory Systems: Casting Application

Sofija Milicic ¹, Amir M. Horr ^{1,*}, Stefanie Elgeti ², Manuel Hofbauer ¹ and Rodrigo Gómez Vázquez ¹

¹ LKR Light Metals Technologies Ranshofen, Austrian Institute of Technology, 1210 Vienna, Austria; sofija.milicic@ait.ac.at (S.M.); manuel.hofbauer@ait.ac.at (M.H.); rodrigo.gomez@ait.ac.at (R.G.V.)

² Institute of Lightweight Design and Structural Biomechanics (ILSB), Vienna University of Technology (TU Wien), 1040 Vienna, Austria; stefanie.elgeti@tuwien.ac.at

* Correspondence: amir.horr@ait.ac.at; Tel.: +43-(0)50550-6918

Abstract

Artificial Intelligence (AI) and Machine Learning (ML) are increasingly driving the digital transformation of manufacturing systems, enabling the transition from conventional process operation toward intelligent, adaptive, and data-centric production environments. This work presents the development of AI-enabled advisory systems for casting processes, integrating singular value decomposition (SVD)-based reduced-order models with a Variational Autoencoder with Arbitrary Conditioning (AC-VAE) and hybrid simulation frameworks to support real-time process prediction and optimization. The proposed approach leverages manufacturing data to establish predictive models capable of rapidly evaluating process conditions, optimizing operating parameters, and enhancing product quality while reducing material waste, energy consumption, and production costs. By combining physics-based understanding with AI-driven analytics, the framework facilitates real-time decision support, adaptive process control, and continuous performance improvement within modern manufacturing ecosystems. These capabilities contribute to the broader objectives of Industry 4.0 and emerging Industry 5.0 paradigms, including automation, connectivity, operational resilience, sustainability, and human-centered manufacturing. A representative Horizontal Direct Chill (HDC) continuous casting case study is presented to demonstrate the practical implementation of the framework, encompassing database generation, model training, validation, and deployment of predictive advisory tools for real-time manufacturing applications.

Keywords: data models; machine learning; neural network; manufacturing processes; digitalization; digital advisory system



Academic Editors: Bo Yang and Jie Zhang

Received: 7 July 2026

Revised: 31 July 2026

Accepted: 14 August 2026

Published: 18 August 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The integration of AI into manufacturing systems, particularly within metal casting operations, represents a significant advancement in the ongoing digital transformation of industry under the Industry 4.0 and Industry 5.0 paradigms. Casting remains one of the most widely utilized manufacturing processes for producing complex metallic components; however, its operation is governed by highly coupled thermo-fluid, metallurgical, and solidification phenomena that span multiple spatial and temporal scales. Product quality is strongly influenced by process parameters such as pouring temperature, casting speed, cooling conditions, and mold design, where even minor deviations can result in defects including porosity, hot tearing, shrinkage, cracking, and residual stress accumulation.

These quality-related issues contribute directly to increased production costs, material losses, and reduced operational efficiency. Conventional process optimization methods, which predominantly rely on experimental trial-and-error approaches, offline numerical simulations, and empirical expertise, are often resource-intensive and lack the flexibility required to support modern manufacturing environments characterized by increasing complexity, customization, and real-time operational demands [1–4].

To overcome these limitations, AI-driven advisory systems have emerged as powerful tools for intelligent process optimization and decision support. By integrating data acquired from experimental studies, sensor networks, and high-fidelity physics-based simulations, these systems establish predictive frameworks capable of capturing the nonlinear, transient, and multi-scale behavior of casting processes, including the evolution of thermal fields, flow dynamics, solidification behavior, and microstructural development. Machine learning algorithms, combined with advanced surrogate modeling and reduced-order modeling techniques, enable the efficient representation of complex process physics while significantly reducing computational requirements. As a result, computationally expensive simulations can be transformed into rapid predictive tools suitable for real-time monitoring, quality assessment, process optimization, and adaptive control [5–9]. Furthermore, these approaches facilitate the generation of large synthetic datasets and expanded process knowledge, supporting the deployment of digital advisory systems, digital shadows, and digital twins that enhance productivity, improve product quality, and accelerate the realization of intelligent and autonomous manufacturing ecosystems.

Furthermore, digital, physics-informed advisory platforms for manufacturing have matured into sophisticated decision-support and increasingly autonomous decision-execution systems that synergistically combine machine learning (ML), statistical inference, optimization algorithms, and physics-based computational models. These frameworks are designed to recommend, or autonomously determine, optimal process parameters to improve critical performance metrics including product quality, production efficiency, operational cost, throughput, and environmental sustainability [10–13]. During the last decade, these technologies have evolved from conventional offline analytical tools and dashboard-oriented monitoring systems toward fully integrated, closed-loop, real-time control ecosystems embedded within digital twin infrastructures, supported by IoT-enabled data acquisition architectures and high-performance computing capabilities. Recent studies across advanced manufacturing sectors have demonstrated a clear transition from descriptive and diagnostic analytics toward predictive and prescriptive intelligence. Nevertheless, several challenges remain unresolved, particularly with respect to industrial-scale data accessibility, data consistency, interoperability between heterogeneous systems, and model explainability. Modern advisory systems increasingly incorporate both forward and inverse predictive functionalities, enabling the modeling, supervision, and optimization of manufacturing operations governed by strongly coupled multi-physical, multiphase, and multiscale interactions, as well as extensive temporal and spatial datasets that exceed the limitations of traditional simulation methodologies [14–16].

This study examines the implementation of digital advisory frameworks for the simulation and design optimization of industrial material-processing operations. The research systematically addresses key technical challenges, including data quality assurance, availability and integrity of industrial datasets, selection of efficient numerical solvers and interpolation methodologies, generation of representative snapshots for reduced-order modeling, validation and verification strategies, and the construction of generative data models together with structured process databases. The principal contribution of this work is the development of an integrated computational framework that combines ML-driven approaches, physics-based simulation tools, and advanced interpolation schemes to

effectively handle both steady-state and transient manufacturing conditions. By incorporating reduced-order modeling techniques and iterative data-centric training strategies, the proposed methodology significantly lowers computational demands while improving predictive robustness and accuracy, particularly for highly complex systems characterized by coupled multi-physical, multiphase, and multiscale behavior. Several industrially relevant case studies are presented to validate the applicability and performance of the proposed framework in real-world manufacturing environments. In addition, the study highlights ongoing advancements in process-data acquisition and synthetic-data generation, while offering practical recommendations regarding solver configuration, interpolation strategy selection, and deployment considerations for industrial implementation.

2. Background and Literature Review

The integration of AI-driven data science techniques into manufacturing processes traces its origins to the early 20th century, beginning with the application of fundamental statistical methods for process control. This evolution has accelerated into a highly sophisticated manufacturing paradigm where data informs product design (e.g., generative models for concept development), process optimization (e.g., predictive algorithms for resource allocation), and process control (e.g., adaptive feedback mechanisms for operational stability). This section provides a chronological overview of this progression, highlighting key milestones, challenges, and technological breakthroughs.

2.1. Historical Background

The historical evolution of manufacturing process modeling is strongly associated with the automation and digitalization advances introduced during the Industry 3.0 era, where programmable logic controllers (PLCs), numerical control (NC) systems, and early robotics established the technological basis for modern intelligent manufacturing systems. Earlier foundations emerged from statistical process control (SPC), which introduced quantitative approaches for monitoring production variability and improving process stability [17–19]. The introduction of the first control chart in 1924 represented a major milestone in data-driven manufacturing, enabling differentiation between common-cause and special-cause process variations and laying the groundwork for statistical quality control methodologies [20–22].

Subsequent developments included Deming's quality management principles, Total Quality Management (TQM), and broader integration of statistical analytics into industrial optimization and decision-making frameworks [23,24]. During the 1950s, the emergence of computational technologies and the formal establishment of artificial intelligence (AI) significantly accelerated manufacturing automation. Early AI applications involved neural networks for pattern recognition and machine learning techniques for defect detection and quality inspection. By the 1960s, NC machine tools incorporated elementary data analytics and optimization methods, while rule-based robotic systems and computer-aided design (CAD) technologies initiated the transition toward intelligent and digitally integrated manufacturing environments [25–28].

During the 1970s, expert systems (ESs) emerged as an important milestone in the evolution of intelligent manufacturing, introducing rule-based artificial intelligence capable of replicating human expertise for complex industrial decision-making. These systems applied engineering knowledge through if–then logic structures to optimize tasks such as process design, equipment configuration, and fault diagnosis, significantly reducing operational errors and manufacturing costs. The integration of fuzzy logic into ES further enhanced their capability to manage uncertainty and nonlinear industrial processes, particularly in chemical and process industries. In the 1980s, a second wave of AI adoption accelerated

the deployment of ES in robotics, production scheduling, and automated planning systems, enabling more adaptive and knowledge-driven manufacturing environments [29–31].

The emergence of Industry 4.0 around 2011 represented a major paradigm shift toward smart and interconnected manufacturing systems. This framework integrates cyber-physical systems (CPS), the Internet of Things (IoT), cloud computing, and advanced AI/ML technologies to support real-time monitoring, predictive maintenance, autonomous control, and process optimization. By combining large-scale data acquisition with computational intelligence, Industry 4.0 established highly connected and data-centric production ecosystems that improve manufacturing flexibility, operational efficiency, and system resilience across global industrial networks [32–39].

The development of surrogate models for engineering design and manufacturing optimization is closely linked to the efficient solution of parametric partial differential equations (PPDEs), where parameters represent geometry, material behavior, boundary conditions, or processing variables [40]. Surrogate modeling approaches can generally be grouped into three categories. The first involves analytical or dimensional reduction techniques that simplify governing equations through physical assumptions, symmetry exploitation, or linearization. The second category includes computational model-order reduction methods such as reduced basis (RB) and proper orthogonal decomposition (POD), which generate low-dimensional representations of high-fidelity numerical simulations by projecting solutions onto dominant modal subspaces. The third category consists of data-driven and physics-informed machine learning approaches, including physics-informed neural networks (PINNs) and Deep Operator Networks (DeepONets), which approximate parametric solution operators using advanced neural architectures.

Among these approaches, physics-informed machine learning has gained significant interest because it combines the flexibility of neural networks with physical conservation laws and governing equations [41–43]. Although earlier concepts originated in the 1990s, the modern PINN framework introduced in 2017 substantially accelerated research and industrial adoption [44]. In manufacturing, PINNs have been applied to solidification prediction, composite curing analysis, and coating-process modeling under limited data conditions [45]. These methods enable physically consistent, computationally efficient, and interpretable surrogate models for multi-physical manufacturing systems requiring rapid prediction and optimization across varying process conditions [46,47].

2.2. Critical Assessment–Process Modelling Developments

Over the past century, manufacturing process modeling has evolved from empirical and statistics-based methodologies into highly integrated, data-driven, and physics-informed intelligent systems. This progression represents one of the most significant technological achievements in industrial history, fundamentally transforming production efficiency, product quality, automation capability, and decision-making. Nevertheless, despite major advances, substantial scientific and industrial challenges remain unresolved.

The earliest developments, particularly SPC and quality engineering methodologies introduced during the early twentieth century, provided the first systematic framework for reducing process variability and improving manufacturing consistency. These approaches were practical, computationally inexpensive, and highly effective for mass production environments. Their implementation in automotive, steel, and electronics industries significantly improved reliability and reduced waste. However, these models were fundamentally limited to descriptive and statistical analysis and lacked predictive capability for highly nonlinear or multi-physical processes.

The computational revolution of the 1950s–1980s introduced numerical control systems, expert systems, and early AI methodologies into manufacturing. These technologies

improved automation, production scheduling, and quality inspection while reducing dependence on manual operation. Expert systems demonstrated strong industrial value in repetitive and rule-based applications, achieving substantial reductions in operational errors and manufacturing costs. However, their rigidity, poor adaptability, and dependence on manually encoded knowledge created major bottlenecks. They struggled with uncertainty, dynamic process behavior, and large-scale multivariable systems typical of industrial manufacturing.

The emergence of Industry 4.0 marked a transformative leap toward interconnected, intelligent, and autonomous production systems. Cyber-physical systems, IoT infrastructures, cloud computing, and machine learning enabled real-time monitoring, predictive maintenance, and adaptive process optimization. Industrial implementation has significantly improved productivity, energy efficiency, supply chain responsiveness, and operational flexibility. Modern manufacturing facilities can now process massive volumes of sensor data and optimize processes continuously under rapidly changing operating conditions.

Despite these achievements, several critical gaps remain. Industrial data are often sparse, noisy, incomplete, or poorly standardized, limiting the reliability of ML-based models. Interoperability between heterogeneous hardware, software, and legacy systems remains a major industrial bottleneck. Furthermore, many AI-driven models operate as “black-box” systems with limited interpretability, creating challenges for industrial trust, certification, safety validation, and regulatory acceptance. High computational cost also remains a barrier for real-time simulation of multi-physical, multiscale, and transient manufacturing phenomena.

Recent developments in surrogate modeling, reduced-order modeling, and physics-informed machine learning represent important attempts to overcome these limitations. PINNs and hybrid modeling frameworks combine physical conservation laws with data-driven learning, improving predictive robustness while reducing data requirements. These developments are particularly important for industrial environments where full experimental measurements are expensive or unavailable.

From a broader perspective, the historical development of process modeling reflects a remarkable human achievement in scientific computing, automation, and industrial engineering. Manufacturing productivity, precision, scalability, and product complexity have increased dramatically over the last century. Modern industries can now produce highly sophisticated products at unprecedented speed, quality, and global scale. Nevertheless, the next major challenge lies not only in increasing model intelligence, but also in achieving reliable, explainable, sustainable, and fully interoperable industrial AI systems capable of autonomous decision-making under real manufacturing constraints. Figure 1 illustrates the pictorial representation of data methods’ background over a century from statistical process control to today’s Industry 4.0 and 5.0 smart factories.



Figure 1. Pictorial representation of data methods’ background over a century from statistical process control to today’s Industry 4.0 and 5.0 smart factories.

3. Data Processing and AI Applications

The advancement of data processing and artificial intelligence in manufacturing has evolved from conventional statistical analysis techniques toward sophisticated cyber-physical production ecosystems that incorporate digital twins, digital shadows, and AI-assisted advisory platforms. With the continuous growth of computational capabilities and industrial data availability, manufacturing systems have progressively transitioned from descriptive analytics to intelligent, data-driven decision-making frameworks. The development of data mining and advanced analytical methods enabled industries to detect hidden correlations, operational inefficiencies, and process anomalies, thereby improving production optimization, quality control, and maintenance strategies. Currently, the integration of AI methodologies with cloud computing, edge computing, and industrial IoT infrastructures supports real-time data acquisition, predictive analytics, and adaptive process control across complex manufacturing environments. The latest stage of this technological evolution is characterized by the deployment of digital twins and digital shadows, which establish continuously synchronized virtual representations of physical assets, processes, and production systems. Combined with AI-enabled advisory and autonomous decision-support systems, these technologies facilitate dynamic process optimization, predictive maintenance, operational resilience, and enhanced manufacturing intelligence under rapidly changing industrial conditions.

3.1. Digital Systems for Manufacturing

The methodological architecture of digital manufacturing systems is founded on a multi-layered framework that integrates data acquisition, preprocessing, modeling, simulation, and intelligent decision-support functionalities.

- The first stage involves comprehensive data acquisition from multiple sources, including experimental measurements, industrial sensor networks, numerical simulations, historical production databases, and data-mining platforms. This ensures extensive visibility into manufacturing conditions, process variables, and equipment behavior.
- Subsequently, the acquired raw data undergoes preprocessing operations such as filtering, normalization, mapping, translation, feature extraction, and interpretation to improve data integrity, consistency, and analytical reliability for downstream applications.
- The next phase focuses on data modeling and advanced analytics, where statistical techniques, dimensionality-reduction approaches, and machine learning algorithms, including regression models, clustering methods, and neural networks, are employed for process optimization, fault diagnosis, anomaly detection, and predictive analysis.
- Building upon these data-driven and physics-informed frameworks, virtualization and simulation technologies enable the implementation of digital twins, digital shadows, and AI-assisted advisory systems. These platforms generate continuously synchronized virtual representations of physical assets and manufacturing operations, enabling real-time monitoring, predictive evaluation, optimization, and scenario analysis under dynamically changing operational conditions.

Finally, intelligent advisory and autonomous control systems utilize advanced AI-based decision engines to deliver prescriptive recommendations or automated corrective actions, supporting the realization of smart manufacturing environments. Overall, the framework emphasizes scalability, interoperability, and cyber-physical integration, allowing industries to transition from reactive manufacturing practices toward predictive and prescriptive operational strategies that enhance productivity, resilience, sustainability, and overall process efficiency. Figure 2 presents a schematic illustration of the methodological framework for digital manufacturing systems.

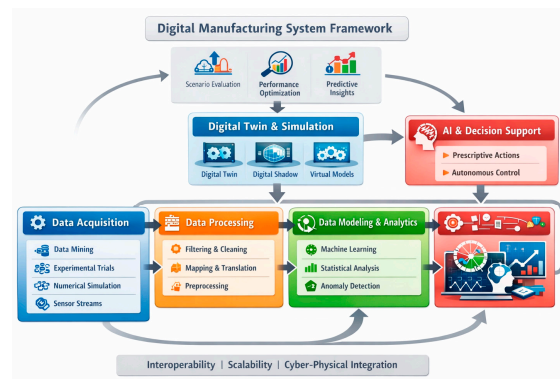


Figure 2. Pictorial representation of methodological framework for digital systems in manufacturing, including data acquisition, processing, modeling, and decision support.

3.2. Data Sources & Databases

A comprehensive framework for establishing data sources and process databases for manufacturing operations, including casting, extrusion, and additive manufacturing (AM), requires structured, interoperable, and semantically consistent data architectures. The inherently coupled multi-physical, multiscale, and multiphase characteristics of these processes, together with highly nonlinear parameter interactions, create major challenges for generating reliable and high-quality datasets suitable for process modeling and intelligent manufacturing applications.

Experimental data source: Experimental data remains the most reliable source of manufacturing information because it is derived directly from physical measurements and real process observations. Such datasets typically provide high accuracy and strong physical credibility; however, experimental campaigns are often expensive, labor-intensive, and time-consuming. For industrial processes with large multidimensional parameter spaces, exhaustive experimentation becomes economically impractical. As a result, experimental datasets are commonly employed for calibration and validation purposes rather than large-scale database generation.

Simulation data source: Numerical simulation data has therefore become a critical component in modern manufacturing database development. Verified and validated simulations enable systematic exploration of process conditions by varying material properties, geometries, and boundary conditions across extensive parameter domains. Advances in computational mechanics, high-performance computing, and multi-physics modeling have significantly enhanced the feasibility of large-scale synthetic data generation for industrial applications.

Mining data source: Data mining approaches provide an additional source of process knowledge by extracting information from published literature, industrial reports, and publicly accessible databases. This strategy is particularly effective for processes characterized by relatively constrained parameter spaces or limited experimental accessibility.

In parallel, live process data acquired through sensors, industrial monitoring systems, and IoT-enabled equipment provides continuous real-time information regarding machine behavior, process dynamics, and operational variability. These dynamic datasets support adaptive learning, feedback control, and continuous model refinement through iterative updating mechanisms. For highly complex manufacturing operations such as casting, extrusion, and AM, robust process databases should integrate validated numerical simulations, targeted experimental measurements, mined datasets, and real-time sensor information within a unified framework. Over time, continuously acquired live process data enhances database accuracy, adaptability, and predictive robustness, thereby supporting advanced applications including digital twins, intelligent process optimization, and

AI-driven advisory systems. Figure 3 presents a schematic illustration of data sourcing methodologies and process database development strategies for manufacturing systems.

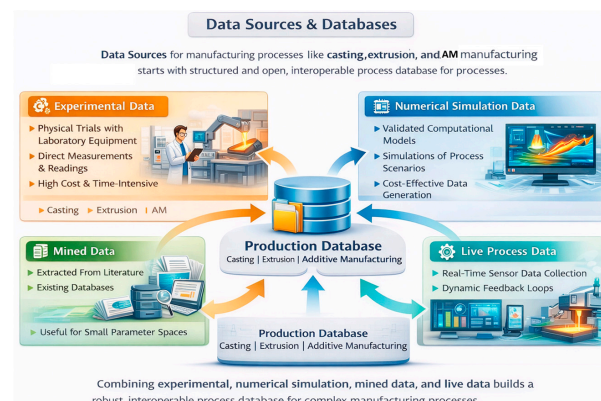


Figure 3. Pictorial representation of methods for data sourcing and process database building using experimental, simulation, mined and live data.

3.3. Reduced-Order Models

Reduced-order models (ROMs) in manufacturing are computationally efficient surrogate modeling approaches designed to capture the dominant physical behavior of complex manufacturing systems while significantly reducing the computational cost associated with high-fidelity simulations. These models are typically derived from physics-based or physics-informed full-order models that describe coupled multi-physical, multiphase, and multiscale phenomena occurring in processes such as casting, extrusion, and AM. ROMs preserve the essential process characteristics, including thermal evolution, fluid flow, material deformation, phase transformation, microstructural development, and non-linear system dynamics, while enabling substantially faster computation compared with conventional finite element or finite volume simulations.

The development of ROMs commonly relies on advanced mathematical and data-driven techniques such as proper orthogonal decomposition (POD), reduced basis (RB) methods, surrogate regression approaches, and machine learning-assisted reduced models. These methods are particularly suitable for industrial manufacturing applications that require rapid or near-real-time predictions across high-dimensional parameter spaces and varying processing conditions.

For manufacturing processes such as casting, a variety of data-solver methodologies can be employed to simplify complex multidimensional parameter interactions and construct efficient reduced representations. These include eigenvalue decomposition techniques, statistical regression models, and classification-based algorithms. To further accelerate prediction capability, ROM frameworks are frequently combined with efficient interpolation techniques such as radial basis functions (RBF) and inverse distance weighting (IDW). In addition, machine learning methods, including neural networks (NNs) and genetic-algorithm-based symbolic regression, can be integrated into the framework to enhance pattern recognition, predictive accuracy, robustness, and adaptive learning capabilities for advanced digital manufacturing applications.

The development of ROMs for manufacturing applications has been extensively investigated in previous studies [48–52], where integrated computational frameworks combining data reduction algorithms, interpolation methods, and machine learning techniques were proposed. In the present work, an eigenvalue-based data-reduction methodology is applied to a multi-physics and multiscale casting process using singular value decomposition (SVD) to decompose and project high-dimensional process data onto a reduced subspace representation. The modeling framework considers four principal process variables: casting speed,

initial melt temperature, cooling-water flow rate, and cooling-water temperature, which collectively define a four-dimensional process parameter domain. The dimensionality reduction introduced by SVD inevitably results in some information loss, as the original high-dimensional solution is represented using a limited number of dominant modes. The objective of the ROM is to preserve the principal physical behavior of the casting process while reducing the complexity of the solution. Consequently, the reduction represents a compromise between computational efficiency and physical fidelity, and its suitability is verified through validation against the corresponding high-fidelity simulations.

Two categories of reduced models are developed. Initially, a compact database containing 16 casting scenarios is generated for verification and validation purposes. These design cases are selected using Latin hypercube sampling (LHS) to ensure efficient and uniform distribution throughout the process parameter space. Following successful model validation, an expanded database of 100 casting conditions is constructed to support large-scale reduced-model development. The same LHS strategy is employed for parameter-space sampling, and the resulting ROM framework is subsequently used to generate approximately 10,000 predictions of thermal and microstructural fields. These datasets provide the training foundation for the digital process advisory system. Figure 4 presents the overall architecture of the reduced-order modeling framework.

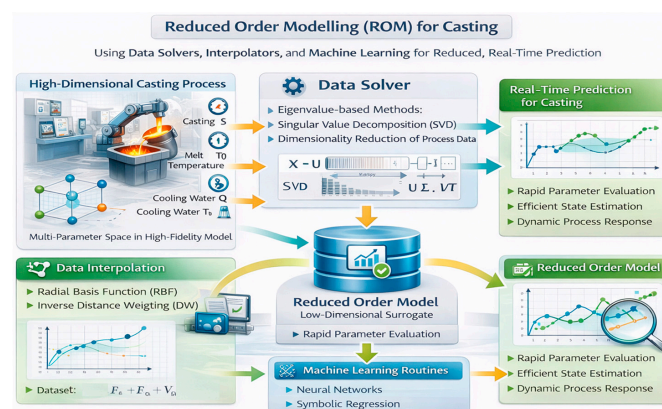


Figure 4. Schematic overview of reduced-order modelling framework including its components: data solvers, interpolators and parameter space.

3.4. Critical Analyses: AI-Based Model Performances

AI-based process modeling has transformed manufacturing from conventional monitoring and offline optimization toward intelligent, adaptive, and autonomous production systems. Among modern digital frameworks, digital shadows, digital twins, and digital advisory systems represent different levels of integration and capability. Digital shadows primarily provide one-way synchronization from physical systems to virtual models, enabling monitoring and historical analysis. Digital twins extend this concept by establishing bidirectional communication between physical and virtual systems, supporting real-time simulation, predictive analysis, and adaptive optimization. Digital advisory systems further integrate AI-driven decision engines that provide prescriptive recommendations or autonomous process control. In terms of performance, digital twins generally offer the highest predictive and operational capability, while digital advisory systems provide superior decision-support functionality for industrial optimization and active process management.

Despite substantial progress, several gaps and bottlenecks continue to limit industrial implementation. The largest challenge remains data quality and availability. Industrial datasets are frequently incomplete, noisy, inconsistent, or non-standardized due to sensor limitations, hardware variability, and fragmented production infrastructures. For highly complex manufacturing operations such as casting, extrusion, and AM, generating suf-

ficiently broad and representative datasets across multidimensional parameter spaces remains expensive and computationally demanding. Additionally, many existing AI models still rely heavily on simulation-generated synthetic data, which may not fully capture real production variability and uncertainty.

Another major limitation involves model interpretability, scalability, and adaptability under changing production conditions. Many machine learning-based systems operate as black-box models with limited physical transparency, reducing industrial trust and complicating validation procedures. Furthermore, process conditions in industrial manufacturing are inherently dynamic due to material fluctuations, equipment aging, environmental variations, and operational disturbances. Static data models often fail to adapt efficiently to continuously evolving process conditions.

To address these issues, modern frameworks increasingly incorporate dynamic databases, live sensor feedback, and active learning strategies. Continuous real-time data acquisition allows AI models and digital twins to update process predictions and improve accuracy through iterative retraining and adaptive optimization. However, implementing reliable closed-loop control systems remains challenging due to cybersecurity concerns, computational cost, interoperability issues, and latency limitations between edge devices, cloud platforms, and production equipment. Therefore, future development of AI-based manufacturing systems requires robust hybrid frameworks that combine physics-based modeling, machine learning, real-time feedback, and interpretable AI to achieve reliable, scalable, and fully autonomous industrial process optimization.

4. AI-Driven Decision Advisory Systems

In the context of manufacturing digitalization, digital advisory systems, often referred to as online process advisory systems, are best understood as AI-enabled design and decision-support tools. These systems can operate atop a process's data infrastructure and its virtual counterparts (such as digital shadows and digital twins), providing guidance for new design scenarios and continuously transforming available process data into real-time predictions and corrective recommendations for engineers and operators. As an AI-powered platform, it leverages real-time predictive models to support process design and execution. It identifies potential failures and quality deviations, recommending optimal parameter sets while enhancing (rather than replacing) human expertise in the manufacturing environment.

4.1. System Architecture

Many industrial manufacturing processes are governed by a set of operational parameters that must be carefully selected, controlled, and continuously monitored throughout production. The appropriate choice of these parameters is critical, as certain combinations may introduce safety hazards, cause unplanned downtime, or compromise product quality. Identifying parameter configurations that ensure stable and reliable operation while simultaneously meeting stringent quality requirements is a complex multidimensional task. In industrial practice, this process often involves iterative trial-and-error procedures and depends heavily on the expertise and tacit knowledge of experienced operators. Moreover, even minor deviations from optimal operating conditions can result in quality deterioration, reduced process efficiency, or production interruptions.

An AI-based advisory system is designed to assist operators by providing data-driven recommendations for feasible and robust process settings under varying operational constraints. In real-world manufacturing environments, not all process parameters can be adjusted freely: some variables can be modified dynamically during operation, whereas others may be temporarily fixed due to physical, technical, or plant-specific limitations.

The advisory task can therefore be formulated as a constrained inverse problem: given a scenario in which a subset of process parameters is fixed, determine appropriate values for the remaining adjustable parameters such that the process operates safely and stably, and the resulting product satisfies predefined quality criteria. By systematically exploring feasible regions of the parameter space, such systems aim to enhance decision-making, reduce reliance on trial-and-error experimentation, and improve overall process robustness and efficiency.

As illustrated in Figure 5, operators first provide the current process state and specify which parameters are fixed. These fixed inputs define the operational constraints within which the AI-based advisory system must identify feasible adjustments. Based on this constrained scenario, a generative model proposes candidate values for the remaining adjustable parameters. Each candidate configuration is subsequently evaluated according to criteria developed in collaboration with domain experts. These criteria quantify process stability, product quality, and any additional process-specific requirements. Configurations that fail to meet the predefined minimum thresholds for any criterion are systematically discarded. The remaining feasible candidates are then ranked using a composite, process-specific performance metric, often referred to as a fitness score, that integrates all relevant quantitative measures into a single evaluation framework. The proposed framework therefore addresses a constrained inverse optimization problem by identifying high-quality feasible process parameter combinations according to the defined fitness score. Rather than guaranteeing a mathematically global optimum, the objective is to efficiently identify high-quality candidate solutions suitable for real-time industrial decision support. The highest-ranking recommendation(s) are ultimately presented to the operators via a dedicated Human–Machine Interface (HMI). With appropriately selected AI architecture and a suitable computational infrastructure, this advisory workflow can be executed within a few seconds, thereby enabling near-real-time decision support during production. While the described forward (online) workflow illustrates the operational use of the system, its overall performance critically depends on the underlying AI model and the robustness of its offline training pipeline. In particular, model selection, data quality, and feature representation significantly influence predictive accuracy and generalization capability. This work focuses on generative AI models and deep learning techniques, which typically require large-scale, high-quality datasets and carefully designed data representations to ensure reliable performance in industrial environments.

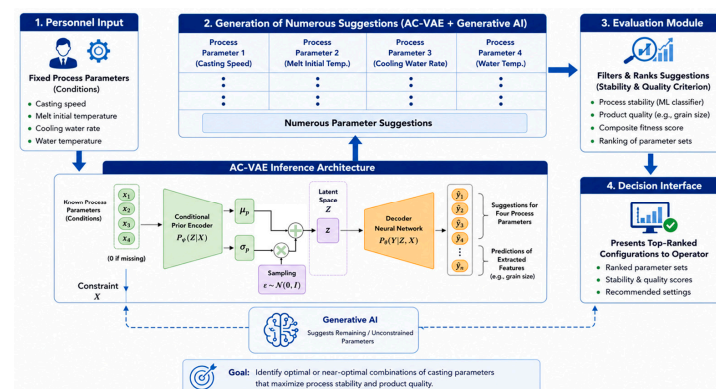


Figure 5. Schematic overview of the proposed AI-based advisory system and its online decision-support workflow.

4.2. Data Infrastructure, Generation & Flow

Ideally, the data acquisition stage should rely on real physical measurements obtained directly from the manufacturing process. In industrial practice, however, such data are often

costly to collect, limited in quantity, and constrained by operational, safety, and production requirements. These limitations restrict systematic exploration of the parameter space and hinder the development of data-intensive AI models. To address this challenge, the present work relies on synthetic data generated through high-fidelity numerical process simulations or data-driven surrogate models derived from them. While high-fidelity simulations provide accurate representations of the underlying physical phenomena, their computational cost can be prohibitive for large-scale data generation. Therefore, complementary strategies for efficient mass data generation are introduced to achieve broad and systematic coverage of the relevant parameter space while maintaining physical consistency.

To enable a robust simulation-driven workflow, a dedicated data infrastructure was developed to ensure reproducibility, traceability, and efficient data management. Process-relevant information is stored in a PostgreSQL database, serving as the central repository for simulation inputs and post-processed outputs. The numerical simulation pipeline is deployed in a cluster computing environment, and simulation runs are orchestrated through automated scripts managing data exchange between the database and the simulation framework. The database schema follows a relational model, whose complexity depends on the number of modeled variables and the level of detail in the simulated process. A central component of the schema is the result table, storing post-processed simulation outputs together with foreign-key references to associated metadata tables, including process parameter combinations, material properties of the modeled system, geometric or mesh-related inputs, and other process- or simulation-specific descriptors. This structured organization ensures full traceability: for any given simulation result, all underlying inputs and configuration details can be retrieved through straightforward database queries, while reproducibility is ensured by persistently storing the complete set of metadata required to rerun a simulation under identical conditions. Consequently, the proposed infrastructure provides a scalable foundation for large-scale synthetic data generation and AI model development in manufacturing applications.

Post-processing constitutes a critical stage, as it determines which components of the raw simulation output are preserved and how they are represented within the database. High-fidelity numerical simulations typically generate three-dimensional fields of multiple physical variables evolving over time on fine-resolution meshes. Depending on the characteristics of the manufacturing process and the definition of the advisory task, either the full transient solution must be retained, or only steady-state information is required. In the latter case, storing solely the final time step is sufficient, and intermediate states can be omitted. Persisting full 3D mesh data for every simulation, particularly for each time step is highly demanding in terms of storage capacity and increases downstream processing complexity. In practice, ML models rarely require complete high-resolution 3D fields, as advisory tasks typically depend on characteristic patterns, physically meaningful summary metrics, or engineered features capturing essential process behavior. Consequently, a central objective of post-processing is the transformation of high-dimensional mesh-based outputs into compact, information-rich representations that preserve the relevant physical content while eliminating redundant spatial detail. This dimensionality reduction step reduces storage requirements, improves computational efficiency, and facilitates model training and inference. During the exploratory development phase, multiple data representation strategies can be evaluated based on compressed versions of the full mesh stored in the database. These may include one-dimensional engineered feature vectors, composed of statistical or physically motivated descriptors such as spatial averages, extrema, gradients, or other characteristic quantities; two-dimensional image-like representations constructed by extracting strategically chosen slices of the three-dimensional; or reduced three-dimensional fields generated by coarsening the original mesh. Multi-modal repre-

representations are also possible, combining tabular process parameters with image-based or reduced-field inputs to exploit complementary information sources. Each representation entails a trade-off between retained physical information, computational cost, and inference time. Different data representations involve different trade-offs between retained physical information, computational cost, and inference time. Richer representations, such as image-based or three-dimensional fields, may capture additional spatial information but require substantially higher computational resources and larger training datasets. For the present advisory application, the selected representation provides a practical balance between predictive accuracy and real-time applicability.

The overarching objective is to identify the simplest representation that enables the advisory system to achieve high predictive performance while meeting real-time inference requirements. Following this exploratory phase, the selected data representation can be fixed and the post-processing pipeline streamlined accordingly, ensuring that only the chosen representation is stored in the database. This consolidation reduces data management complexity, enhances scalability, and allows subsequent efforts to focus on expanding the dataset and refining the AI model architecture.

4.3. Mass Data Generation

This study presents a systematic multi-stage data modeling framework developed to address one of the primary challenges in intelligent manufacturing systems: the limited availability of high-quality, large-scale, and uniformly distributed process datasets. The proposed methodology enables the transformation of a relatively small but high-fidelity dataset, derived from experimental measurements and validated computational fluid dynamics (CFD) simulations, into an extensive synthetic process database containing more than 10,000 manufacturing scenarios. This enriched dataset is subsequently employed for training computationally intensive machine learning models within a real-time digital process advisory environment.

The framework is initiated through the construction of reduced-order models (ROMs) based on eigenvalue-driven decomposition techniques combined with inverse distance weighting (IDW) interpolation methods. These initial surrogate models are calibrated and validated against high-fidelity CFD simulations to ensure predictive reliability and physical consistency. The validation strategy incorporates additional design-of-experiments (DOE) cases covering standard operating conditions, near-threshold regimes, and extreme process scenarios in order to evaluate model robustness across the multidimensional parameter domain. Following successful verification and validation, an expanded sampling methodology is implemented to generate a larger snapshot database consisting of approximately 100 casting simulations. These CFD-generated datasets establish an enriched response database spanning critical process variables, including casting speed, melt temperature, cooling-water flow conditions, and thermal boundary parameters. The resulting high-dimensional database forms the foundation for large-scale synthetic data generation, adaptive learning, and intelligent process optimization within advanced manufacturing advisory systems.

Building upon the validated response database, a scalable framework for large-scale synthetic data generation is developed to support AI-driven manufacturing applications. A second-generation ROM is constructed using SVD-based data solvers coupled with IDW interpolation techniques, enabling efficient mapping of process responses throughout the multidimensional parameter space. Using this framework, a large snapshot matrix containing more than 10,000 parameter combinations is generated, and the corresponding synthetic process outputs are rapidly predicted using the ROM rather than computationally expensive CFD simulations.

Because the reduced-order framework preserves the dominant eigenmodes and principal physical characteristics of the original high-fidelity system, the generated predictions remain physically meaningful, thermally consistent, and smoothly distributed within the calibrated operating domain while requiring only a small fraction of the computational resources associated with full-order numerical simulations. Consequently, the proposed methodology enables substantial reductions in computational cost and simulation time without significantly compromising predictive accuracy.

The ROM-generated synthetic database retains the fundamental structural relationships, physical constraints, and process correlations embedded within the validated CFD datasets. This enriched dataset is subsequently incorporated into the training workflow of deep-learning-based predictive models, including neural-network architectures that require extensive and representative datasets for robust learning and generalization. By utilizing ROM-assisted synthetic data generation, the framework overcomes the major limitations associated with large-scale experimental campaigns and repeated high-fidelity simulations, which are often economically and computationally prohibitive.

Figure 6 presents the overall workflow for large-scale synthetic data generation and its integration into the AI-based training architecture for online manufacturing advisory and real-time process optimization systems.

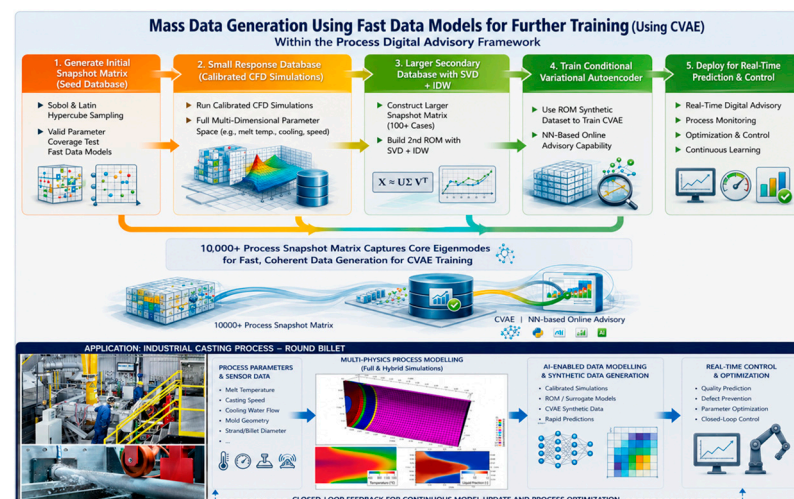


Figure 6. Workflow for the generation of mass synthetic data for further data training using the data model.

4.4. Neural Network Learning

The selection of an appropriate neural network architecture is strongly dependent on the underlying data representation and structural characteristics of the manufacturing dataset. For one-dimensional tabular data, feed-forward neural networks (NNs) are generally the most suitable due to their effectiveness in handling structured numerical feature vectors. Convolutional neural networks (CNNs) are better suited for two-dimensional image-based representations, while graph neural networks and other specialized architectures are often preferred for three-dimensional, spatially connected, or irregularly structured datasets. The present study primarily focuses on tabular process representations while also exploring the potential advantages of multimodal learning frameworks that integrate complementary image-based process information.

Within the tabular-data framework, each simulated process scenario is represented by an n -dimensional feature vector containing quantitative descriptors of process parameters and key physical response variables extracted from numerical simulation outputs. Under this formulation, the process advisory problem can be interpreted as a feature imputation

task, where a subset of process parameters is predefined and the objective is to predict the remaining unknown variables, including adjustable operating conditions and associated physical responses.

To address this challenge, generative AI approaches provide a principled framework. In this work, Variational Autoencoder (VAE)-based architectures are investigated [53]. VAEs belong to a class of probabilistic generative models that learn compact latent-space representations, enabling generation of physically plausible data samples. The Conditional Variational Autoencoder (CVAE) extends this formulation by incorporating conditioning variables that guide the generation process according to specified process constraints; detailed mathematical formulations of VAEs and CVAEs are provided in [54]. Within the context of process modeling and feature imputation, CVAEs enable reconstruction of missing or masked process variables by exploiting learned latent representations conditioned on available process information. However, the standard CVAE formulation assumes a fixed conditioning structure established during the training stage, which may reduce flexibility under dynamically changing manufacturing conditions.

In real industrial advisory scenarios, however, this assumption is overly restrictive. Different operating situations may involve different subsets of fixed parameters; in principle, any subset of process variables may be observed or unknown at inference time. To address this limitation, this work adopts the Variational Autoencoder with Arbitrary Conditioning (AC-VAE) [55], which is trained to handle arbitrary subsets of observed features and infer the remaining components. From a probabilistic standpoint, this flexibility arises from a crucial modification of the conditional prior distribution. In a standard CVAE, the latent variable Z is assumed to follow a standard normal prior independent of the conditioning variables. In contrast, AC-VAE models the conditional prior $P(Z|X)$ as a Gaussian distribution:

$$P(Z|X) = \mathcal{N}(\mu_P(X), \Sigma_P(X)), \quad (1)$$

where the mean $\mu_P(X)$ and covariance $\Sigma_P(X)$ are learned functions of the available conditioning variables X . Architecturally, while a CVAE consists of two NNs, the encoder and the decoder, an AC-VAE employs three networks: (i) the proposal encoder Q (ii) the decoder P ; and (iii) a conditional prior encoder approximating $P(Z|X)$. The complete training architecture is illustrated in Figure 7.

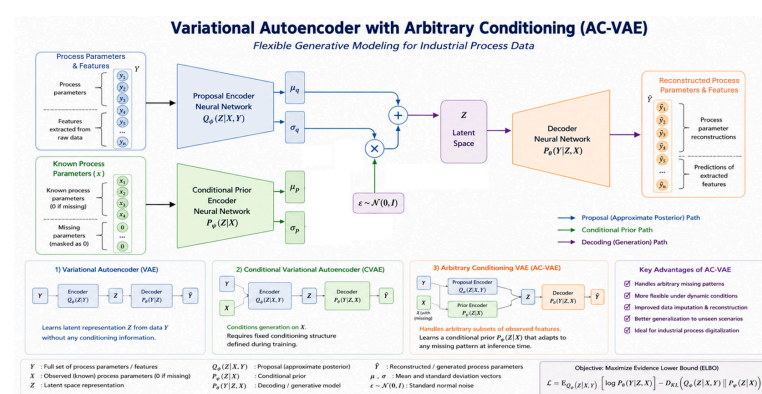


Figure 7. Training-time architecture of the AC-VAE model.

The training of NNs has become highly accessible through modern high-level libraries. In this work, the custom AC-VAE architecture was implemented in PyTorch (v2.8.0), selected for its flexibility and suitability for non-standard models. To reduce boilerplate code and streamline experimentation, PyTorch Lightning was employed, enabling structured training workflows and seamless integration of advanced features such as multi-GPU training, experiment logging, and automated hyperparameter optimization. Optimization is

based on gradient descent, enabled by the differentiability of the network architecture and its loss function. The objective combines two terms: a Mean Squared Error (MSE) term that measures reconstruction fidelity under conditioning constraints and a Kullback–Leibler Divergence (KLD) term regularizing the latent space by aligning the approximate posterior with a conditional Gaussian prior. The closed-form expression of the KLD under Gaussian assumptions ensures stable and efficient training. The Adam optimizer was chosen for its robustness and adaptive gradient handling, and additional regularization techniques were explored to mitigate overfitting.

The dataset is partitioned into training, validation, and test sets using a 60:20:20 split, with Z-score standardization applied to harmonize feature scales. To handle partially observed inputs, each sample is paired with binary masks encoding observed and unknown parameters. The constraint vector is concatenated with its mask, enabling the model to distinguish masked entries from genuine zero-valued features. In the bimodal setting, cross-sectional mesh data are interpolated into 2D image representations and processed through a U-Net-inspired convolutional encoder [56] without “skip connections”. Extracted image features are fused with tabular features and passed to the AC-VAE. Reconstruction is performed symmetrically, with an additional pixel-wise MSE loss incorporated to enable end-to-end optimization of the combined CNN–NN architecture.

At inference time, the proposal encoder is discarded, and only the conditional prior encoder and decoder are retained for generation (Figure 5). While standard regression metrics can be computed on test data, evaluating the true generative capacity remains challenging. Although generated samples satisfy imposed constraints and are drawn from the learned data distribution, their physical consistency cannot be verified without coupling to external simulations (e.g., a ROM). Consequently, the present study evaluates generative performance only in the restricted setting of fully observed test cases using regression metrics, with the coefficient of determination (R^2) preferred for its normalized and interpretable scale. This restricted evaluation represents a limitation of the proposed generative AI framework. Addressing this limitation by coupling the model with external ROM simulations constitutes an important direction for future work, enabling more meaningful assessment of the model’s generative capabilities.

4.5. Process Modelling Applications

AI-based digital advisory systems are increasingly being deployed across modern manufacturing ecosystems to enable intelligent process design, adaptive optimization, and real-time process control. In metal casting processes, these predictive systems are extensively used to optimize melt flow behavior, thermal gradients, solidification kinetics, and defect formation mechanisms. By integrating CFD simulations, thermal sensing data, and machine learning-based surrogate models, digital advisory systems can predict critical casting defects such as shrinkage porosity, hot tearing, macro-segregation, and inclusions before production execution. For example, in continuous and horizontal casting operations, AI-assisted frameworks can dynamically correlate process parameters, including casting speed, cooling intensity, mold geometry, and alloy composition, with local thermal histories and grain growth evolution.

In metal extrusion and forming operations, AI-driven predictive systems are utilized to monitor strain evolution, die-material interactions, recrystallization phenomena, and residual stress development. During hot extrusion of aluminum or magnesium alloys, for instance, digital advisory systems can predict local strain-rate distributions, temperature-dependent flow stress behavior, and dynamic recrystallization kinetics. Such predictive capabilities allow optimization of ram speed, billet temperature, lubrication conditions, and die design to minimize defects such as surface cracking, central burst, and non-uniform

grain structures. Similarly, in forging and sheet-forming applications, AI-assisted digital advisory integrates off-line numerical simulations with in-line sensor measurements (e.g., load cells, thermocouples, infrared imaging, and acoustic emission sensors) to continuously estimate forming loads, die wear progression, and material flow characteristics.

AM represents another major application domain for AI-enabled digital advisory systems due to the highly transient and localized nature of layer-wise material deposition processes. In directed energy deposition (DED), predictive AI frameworks are used to model melt pool dynamics, torch-material interaction, thermal accumulation, and phase transformation behavior. By combining high-speed imaging, pyrometry data, and multi-physics simulations, digital advisory systems can estimate melt pool geometry, cooling rates, porosity-formation probability, and residual stress evolution in real time. Furthermore, AI-assisted microstructure prediction models can estimate grain morphology, texture evolution, and precipitation behavior across multiple deposited layers, enabling the production of components with tailored mechanical properties and improved structural reliability.

5. Case Study: Casting Application

To demonstrate the practical applicability of the proposed digitalization framework, a representative casting use case is introduced. This demonstrator focuses on steady-state casting process conditions, deliberately excluding transient effects associated with process start-up in order to ensure a controlled and reproducible evaluation environment. By concentrating on stable operating regimes, the case study enables a clear assessment of how data-driven models can support process design and decision-making under realistic production conditions. The implementation integrates the essential components of a modern digital workflow, including systematic data acquisition, preprocessing, model development, and training based on an AC-VAE. Within this framework, an online advisory system is realized that provides real-time insights and actionable recommendations.

5.1. Experimental & Simulation Setup

Horizontal Direct Chill (HDC) casting is a continuous industrial process widely employed for the production of high-grade aluminum billets. In this process, molten aluminum undergoes initial heat extraction while flowing through the mold (primary cooling). Upon exiting the mold, the majority of the remaining thermal energy is removed by a circumferential jet of chilled water (secondary cooling). The high cooling rates achieved in HDC casting are critical for establishing the microstructural characteristics and mechanical properties required for subsequent downstream operations, most notably extrusion.

Stable operation of the HDC process depends on precise regulation of key process inputs. Persistent deviations from nominal operating conditions can rapidly impair billet quality, lead to unplanned production interruptions, and potentially introduce safety hazards. In the present study, billets of A6082 aluminum alloy were cast under industrial conditions. Process parameters were actively controlled until nominal steady-state conditions were attained. The internal thermal evolution of the billet was monitored using thermocouples inserted into the melt at predefined locations within the mold. These sensors traveled with the advancing billet, enabling the acquisition of temperature histories along multiple axial trajectories. This measurement strategy further allowed reconstruction of the solidification sump geometry.

A high-fidelity numerical model of the process was developed within the OpenFOAM computational fluid dynamics (CFD) framework. The model employs an Eulerian thermo-fluid formulation, conceptually aligned with the approach proposed by Bennon [57] and further extended to incorporate non-equilibrium solidification [58] effects and microstructural growth kinetics [59]. Assuming lateral symmetry, the computational domain represents

one half of the billet and is discretized using a finite-volume mesh (the supporting files can be found in: <https://gitlab.com/ait-lkr/directchillcasting>). The mesh remains fixed throughout the simulation, with designated inlet and outlet boundaries at the axial extremities. Custom boundary conditions are implemented at the external walls to represent heat extraction by the mold (primary cooling) and by the chilled water jets (secondary cooling). The CFD implementation used for these simulations is provided 774 as open-source code in the Supplementary Materials.

Key geometrical and operational inputs are parameterized to facilitate systematic database generation. Among these, the initial melt temperature, casting speed, chill-water flow rate, and chill-water temperature were selected for variation in order to quantify their influence on process outcomes. Once quasi-steady conditions are reached, the solution fields are time-averaged over multiple time steps. The results are subsequently post-processed to extract thermal histories along selected axial paths, together with predicted microstructural characteristics at the billet outlet plane. Figure 8 illustrates the overlap of the simulation domain with the actual HDC process, including its volumetric 3D discretization and melt flow streamlines mapped with thermal results.

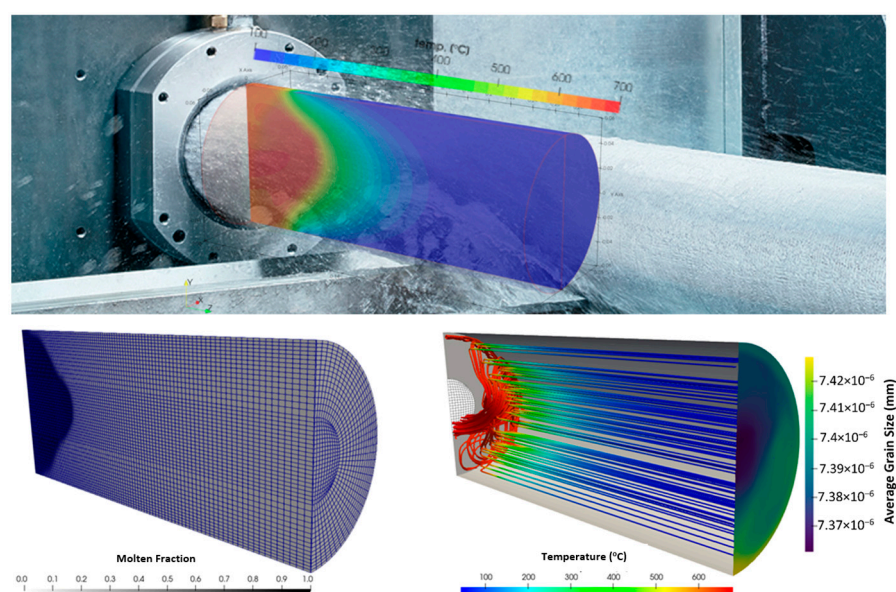


Figure 8. (top) Symbolic overlap of the simulation domain showing thermal results with the actual HDC process. (bottom-left) Volumetric domain discretization. (bottom-right) Melt flow streamlines mapped with thermal results, along with eventual average grain size results at the outlet domain boundary.

5.2. Database Building

The development of reliable process digital advisory systems depends fundamentally on constructing accurate, balanced, and well-covered databases that represent the full multidimensional parameter space of a manufacturing process. An effective strategy follows a tiered, progressive approach, beginning with a small, high-fidelity seed database for model validation and then expanding toward large synthetic databases generated through ROM-based process scenarios. The initial seed database is generated using Sobol or LHS techniques to ensure uniform, balanced, and non-clustered coverage of influential process parameters. For the present work on casting processes, these parameters include melt temperature, cooling rate, casting speed, water-cooling temperature, and other domain-specific variables. Each seed scenario is simulated using calibrated and experimentally validated CFD models, ensuring that the resulting response database faithfully captures the governing thermal and physical behavior of the process. This compact but well-distributed dataset provides the foundation for validating the initial data models, including their

decomposition (e.g., POD/SVD) and interpolation strategies, which together form the mathematical backbone of the physics-informed data modeling framework.

Once the seed models demonstrate sufficient accuracy, a larger secondary snapshot matrix, typically comprising around 100 CFD-based process scenarios, is constructed to span the multidimensional parameter domain with greater resolution. Table 1 presents representative process scenarios with varying operating parameters considered in the CFD simulations. This expanded dataset enhances the development of advanced surrogate and ROMs by capturing a broader range of process variability while maintaining computational efficiency. At this stage, ROM predictions are rigorously validated against independent DOE cases, encompassing nominal, near-boundary, and extrapolated operating conditions. This validation framework ensures model robustness and predictive reliability across the entire process space before further scale-up and deployment.

Table 1. Process scenarios with four varying process parameters for CFD simulations.

Scen_Index	Melt Temp. [°C]	Casting Speed [mm/s]	Water Flow [L/s]	Water Temp. [°C]	Scen. Status ¹
1	713	2.8	2.8	21	Near-Limit
2	720	1.6	2.9	30	Normal
3	708	1.4	1.5	21	Normal
4	745	2.8	1.2	18	Extreme
5	681	0.8	3	17	Near-Limit
6	718	1.8	0.5	17	Normal
7	656	2.3	1.5	21	Normal
8	712	2.6	0.6	21	Normal
9	724	1.8	1.3	23	Normal
10	666	1.6	1.4	19	Normal
...	...	...	...	...	...
98	710	1.9	2.6	28	Normal
99	701	2.1	2.7	29	Normal
100	739	2.1	1.1	29	Near-Limit

¹ Process conditions (normal, near limit, extreme).

With a validated ROM established, the methodology advances to the generation of very large synthetic datasets (10,000+ scenarios). Using reduced-order models that encode the dominant eigenmodes of the physical system, these synthetic datasets can be produced extremely rapidly (on the order of 900 s) while maintaining physical coherence, thermal consistency, and smooth behavior across parameter variations. Because the ROM inherits its structure from physics-validated CFD simulations, every predicted scenario remains grounded in the true physical constraints of the process, even when exploring large and dense parameter grids. These mass synthetic datasets then serve as the training foundation for the AC-VAE module within the digital advisory framework, enabling real-time predictive–corrective process design insights and intelligent online process-parameter recommendation capabilities.

5.3. Variational Autoencoder with Arbitrary Conditioning

Extensive exploratory data analysis (EDA) was conducted on the ROM dataset to systematically assess the influence of process parameters on the raw output data. Based on these findings, feature engineering was performed to effectively represent each case as a high-dimensional feature vector. Derived features included melt pool depth, statistical descriptors of the grain size distribution at the outlet cross-section, temperature values sampled along the billet’s central axis and surface at key locations, and multiple indicators related to liquid fraction evolution. To mitigate redundancy and multicollinear-

ity, highly correlated variables were eliminated through feature selection, resulting in a compact representation of nine properties: four primary process parameters and five engineered features. For the bimodal configuration combining tabular and image-based data, the final billet cross-section was selected as the principal image representation due to its direct relevance to product quality. Scalar mesh values were projected onto a regular pixel grid, and linear interpolation computed pixel intensities as distance-weighted averages of neighboring mesh values, yielding smooth, physically consistent images. The inclusion of a longitudinal section as complementary image input remains a subject for further investigation.

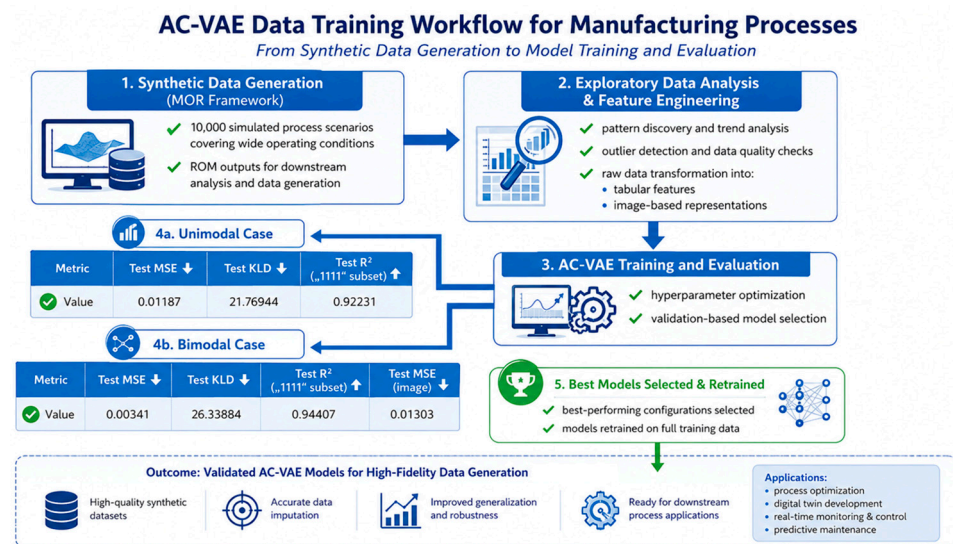
The hyperparameter (HP) search space encompassed both architectural and training-related variables. Architectural hyperparameters included the number of hidden layers in the encoders and decoder of the AC-VAE, latent space dimensionality, and activation functions; in the bimodal setting, the convolutional encoder depth was also varied. Training-related hyperparameters, batch size, learning rate, weight decay, and the use of batch normalization, were systematically explored. The training objective combined tabular MSE and KLD (typically down-weighted) with an additional pixel-wise MSE term in the bimodal case; relative weights of these components were treated as hyperparameters and varied across trials. To ensure comparability between configurations, model selection was based exclusively on the tabular validation MSE. Early stopping was employed to mitigate overfitting. In total, 500 hyperparameter configurations were evaluated independently for the unimodal and bimodal settings, and the best models were selected based on validation loss.

The workflow summarized in Figure 9 outlines the principal stages of the proposed methodology. After completing hyperparameter optimization for both AC-VAE configurations, the best-performing configurations were retrained. Their performance was evaluated on an independent test set, with metrics reported in Figure 9. Both configurations demonstrate strong generalization capability, reflected by low reconstruction MSE values on the tabular features. Although the KLD is reported for completeness, its absolute magnitude is not directly comparable across models due to differences in latent space dimensionality. Nonetheless, the observed KLD values indicate an appropriate balance between latent space regularization and reconstruction fidelity. In particular, the magnitude of the KLD suggests active utilization of the latent variables, avoiding posterior collapse while maintaining high reconstruction accuracy.

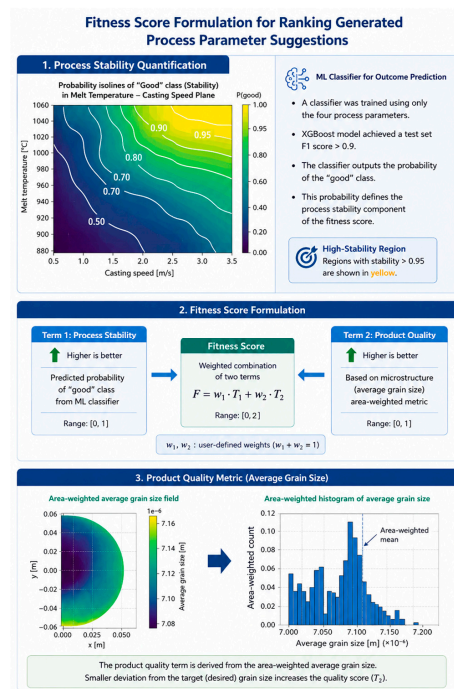
A notable reduction in tabular reconstruction MSE is observed in the bimodal configuration, demonstrating that the inclusion of image-based information enhances tabular reconstruction. However, a fundamental limitation remains the absence of direct metrics for assessing generative capacity at inference time without coupling the model to online physical simulations. In the restricted scenario where all four process parameters are known (the “1111” subset in Figure 9a), generative performance is evaluated using the R^2 . The global R^2 , averaged across features, shows a slight improvement for the bimodal model. More pronounced gains are observed on a per-feature basis, particularly for grain-size-related quantities, attributed to the explicit integration of grain-size information through the image modality during training. Despite these advantages, image reconstruction quality remains limited. Although the per-pixel MSE appears low, it is computed on scaled data and pixel-wise; qualitative inspection reveals noticeable reconstruction errors.

For integration into the proposed advisory system (Figure 5), the AC-VAE inference architecture generates multiple candidate parameter configurations under specified constraints. These suggestions are quantitatively evaluated with respect to process stability and product quality. A composite fitness score was defined to enable systematic filtering and ranking of generated candidates. Process stability is defined as the likelihood that a suggested parameter set produces acceptable outcomes while avoiding bleed-out and

freeze conditions. While a deterministic stability assessment exists within the high-fidelity CFD framework, its reliability diminishes after the data reduction required for ROM construction, as full-field mesh information is no longer retained. To overcome this limitation, a ML classifier was trained to predict outcome classes using only the four process parameters. An XGBoost model achieved strong predictive performance, with a test set F1 score exceeding 0.9. The classifier's probabilistic output defines the stability component of the fitness score as the predicted probability of the “good” class. Figure 9b illustrates probability isolines in the melt temperature–casting speed plane, highlighting regions where stability exceeds 0.95.



(a)



(b)

Figure 9. (a) Overview of the proposed ML workflow and key results; (b) illustration of the fitness score formulation used for ranking generated process parameter suggestions.

The product quality component of the fitness score is defined as the fraction of the predicted grain-size distribution falling within two predefined thresholds, which is to be

maximized. In the unimodal configuration, these thresholds and corresponding quality features must be specified prior to model training. By contrast, the bimodal configuration provides greater flexibility, as quality criteria (e.g., imposing an upper bound on maximum grain size) can be modified dynamically through adjustments to the fitness function without retraining the model. In both unimodal and bimodal settings, generating 100 candidate solutions prior to filtering requires approximately 0.5 s on average, satisfying real-time operational requirements. Since candidate generation is inherently parallelizable, this runtime can be further reduced through parallel execution. Table 2 summarizes the results.

Table 2. Test-set performance and inference time for unimodal and bimodal AC-VAE configurations.

	MSE (Tabular)	KLD	MSE (Image)	R^2 ("1111" Subset)	Inference Time (100 Candidates Generated)
unimodal	0.01187	21.76944	-	0.92231	0.5 s
bimodal	0.00341	26.33884	0.01303	0.94407	0.5 s

5.4. Performance Analyses: Casting Application

To validate the performance of process modeling within the digital advisory framework, comprehensive evaluations of predictive performance were conducted. These assessments focused on accuracy, robustness, and validation protocols (e.g., defined error thresholds) to ensure reliability under real manufacturing conditions. Particular attention was given to operational resilience under varying process environments, as well as compliance with established industrial standards. For the predictive analyses presented in this study, advanced computational tools, including CFD and SVD solvers, robust interpolation schemes, and AC-VAE-based training architectures, were systematically employed. These approaches enabled the development of efficient and scalable predictive frameworks.

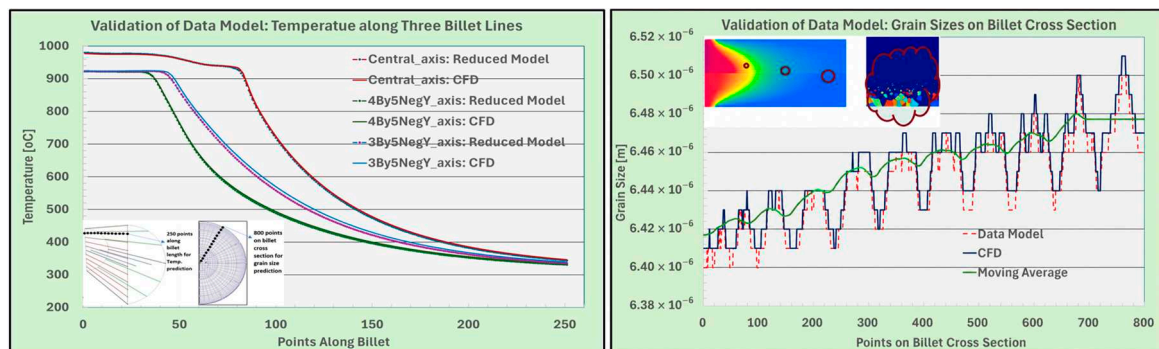
Temperature and Microstructure Predictions: The performance of data-driven models for horizontal casting processes was evaluated to assess their predictive accuracy and their ability to generate reliable large-scale datasets for subsequent model training. Initially, a set of DOE-based CFD simulations was conducted to validate the data models. These simulations covered a broad range of process parameters and operating conditions, including nominal regimes, near-operational limits (e.g., casting machine constraints), and selected extreme scenarios.

Table 3 presents the DOE process scenarios used for validating the data-driven models, encompassing nominal, near-limit, and extreme operating conditions. Figure 10 shows the predicted temperature and microstructure distributions along the longitudinal direction of the cast billet, together with radial profiles across the cross-section of a round billet. The results compare CFD-based DOE simulations with the corresponding predictions generated by the data models. The performance assessment demonstrates that the trained model achieves good accuracy in predicting both the thermal field and microstructural evolution. Minor deviations from the DOE results are observed within the melt pool region, where relatively low temperature gradients create flatter profile characteristics. Nevertheless, the overall agreement between the CFD simulations and model predictions confirms the capability of the data-driven approach to reliably capture the key thermal and metallurgical features across the investigated process conditions. In contrast, within the solidification zone, where steep thermal gradients occur, prediction errors increase (up to approximately 2.5%). This reflects the inherent challenge of capturing rapid thermal variations in data-driven frameworks.

Table 3. DOE scenarios with four varying process parameters for validation of data models.

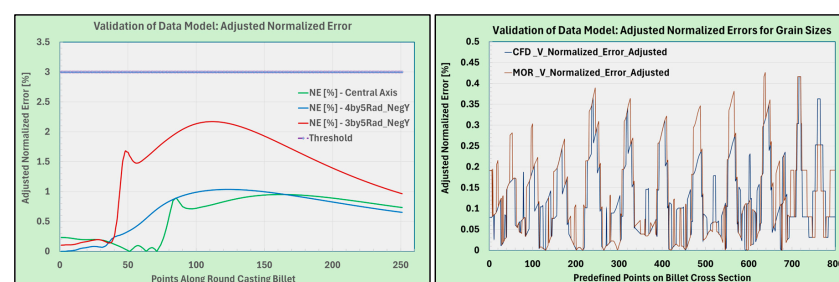
Scen_Index	Melt Temp. [°C]	Casting Speed [mm/s]	Water Flow [L/s]	Water Temp. [°C]	Scen. Status ¹
V1	704	2.9	0.7	17	Near-limit
V2	651	2.5	0.8	16	Normal
V3	699	2.6	2.5	27	Normal
V4	728	2.6	2.5	27	Normal
V5	672	0.6	3	28	Near-limit
V6	748	1.7	2.9	16	Extreme

¹ Process conditions (normal, near limit, extreme).

**Figure 10.** Data model performances for thermal and microstructure predictions against CFD simulations for some DOE cases.

For microstructural predictions, which are inherently discontinuous at specific spatial points, a moving-average filter with a broad kernel (e.g., 50 data points) was applied. This smoothing approach yields continuous grain growth profiles that are more suitable for practical engineering analysis and interpretation.

Error Estimation and Validation: The preliminary thermal and microstructural results indicate that, although the data-driven models show strong potential for process prediction within the digital advisory framework, a more stringent and comprehensive validation scheme is required. Given the complexity of real industrial processes, characterized by high-dimensional parameter spaces, variable operating conditions, and multi-scale grain evolution phenomena, the digital advisory system must reliably capture and predict final process outcomes with high fidelity. Figure 11 presents normalized and adjusted error distributions for both thermal and microstructure predictions. These errors are evaluated along streamlines in the cast billet and along radial profiles within the round billet cross-section. A predefined error threshold of 3% is indicated, reflecting practical industrial requirements and protocol compliance.

**Figure 11.** Data models adjusted normalized errors for thermal and microstructure predictions against CFD simulations for some DOE cases.

The error analysis was performed using predictions from the validated data models, with results normalized and calibrated against the corresponding CFD-based DOE simulations. As anticipated, the largest temperature prediction errors are observed within the solidification zone, where thermal gradients are most pronounced, specifically in the regions corresponding to points 80 to 130 in the error profiles. This behavior highlights the challenges associated with accurately capturing steep thermal gradients and rapid phase transformation dynamics in data-driven modeling approaches.

6. Discussions and Challenges

The deployment of reduced-order and physics-informed, data-driven process models within digital advisory systems, digital shadows, and digital twins is hindered by several fundamental and interrelated challenges. Foremost among these are data availability and data quality, as many manufacturing processes lack sufficiently rich, well-distributed, and high-resolution datasets to span their full multidimensional parameter spaces. Sparse, noisy, or inconsistent process data undermine model robustness, predictive accuracy, and generalizability, limitations that become especially critical for transient, multi-physical, or multi-scale processes. Furthermore, achieving adequate and balanced coverage of the parameter space remains difficult. Even when data is available, it is often unevenly distributed across normal, near-boundary, and extreme operating regions, leaving real-time advisory systems vulnerable to “blind spots” and degraded predictive performance when confronted with previously unseen conditions.

A further challenge lies in data model fitting, as reduced-order and physics-informed models frequently struggle to capture steep gradients, phase transitions, abrupt nonlinear behaviors, and complex initial or boundary conditions. These issues are particularly pronounced in processes characterized by rapid thermal fluctuations, solidification fronts, or large deformation zones. Compounding these difficulties are rate-dependency effects, since many manufacturing operations (including vertical, horizontal, and high-pressure die-casting processes) exhibit rapid heating and cooling cycles, or repetitive cyclic behavior, that often lead to time-shifted, unstable, or inaccurate ROM predictions. Together, these challenges highlight the inherent complexity of integrating high-fidelity data-driven models into advanced process digitalization frameworks.

The integrating ROM-based predictive models into digital process advisory framework adds an additional layer of difficulty. These frameworks require near-instantaneous predictions, high reliability across the entire operational window, and compatibility with heterogeneous data sources including experimental measurements, CFD/FE simulations, and live sensor streams. Ensuring semantic consistency through ontologies, enabling continuous model updating, and maintaining physical validity while generating large synthetic datasets (e.g., via ROM or AC-VAE) are all essential yet technically demanding. Within the context of the proposed framework, these challenges define the practical applicability boundaries of the presented methodology. The advisory system is intended for operation within the process domain represented by the offline-generated training data and should not be expected to extrapolate reliably to previously unseen operating conditions. Compared with conventional industrial practice, which often relies on operator experience, predefined operating windows, or computationally intensive simulation studies, the proposed approach enables rapid generation and evaluation of multiple feasible process parameter combinations suitable for near real-time decision support. Nevertheless, the overall prediction performance remains dependent on the quality and representativeness of the training data, while the reported prediction errors reflect the combined effects of ROM approximation, data representation, and the generative AI model.

These aspects should therefore be considered when deploying the proposed framework in industrial environments.

7. Concluding Remarks

The results indicate that physics-informed, reduced-order data models have the potential to substantially speed up process design, optimization, and decision-making in manufacturing. However, their effectiveness depends on a robust data strategy in which database creation, validation, and continuous refinement are treated as essential engineering tasks. Starting from calibrated simulations and a limited number of high-quality experimental measurements, compact and well-balanced snapshot matrices, constructed using methods such as Sobol or Latin Hypercube sampling, can provide a reliable basis for eigen-based decomposition approaches, including SVD and POD, as well as for the informed selection of suitable solver–interpolator combinations. After validation under normal operating conditions, near-boundary cases, and extreme DOE scenarios, with thermal and microstructural prediction errors remaining below a 3% threshold consistent with practical industrial requirements, these models can be confidently expanded to generate large, physically consistent synthetic datasets. Such datasets can then support advanced learning frameworks, including AC-VAEs, and enable real-time advisory capabilities.

At the same time, the study emphasizes several critical challenges that must be resolved before these models can be reliably integrated into digital shadows and digital twins. Key requirements include ensuring sufficient data availability and quality, achieving balanced coverage of the relevant parameter space, and accurately representing initial and boundary conditions. Further challenges arise from the need to capture steep thermal and mechanical gradients, phase transformations, rate-dependent transient behavior, and multi-scale, multi-physics interactions. To support robust deployment, these complexities must be embedded within database architectures that are semantically consistent, adaptable, and capable of evolving over time.

These studies also demonstrate the considerable potential of combining ROM-based large-scale synthetic data generation with generative machine learning approaches such as AC-VAEs. Such hybrid workflows make it possible to rapidly produce tens of thousands of physically consistent process scenarios, thereby supporting inverse design, uncertainty quantification, and fast advisory inference while requiring only a fraction of the computational resources associated with full CFD or finite element simulations. Nevertheless, the computational acceleration achieved through these approaches introduces additional challenges. The models must remain capable of handling steep data gradients, sudden phase transitions, rate-dependent variations, changing numerical domains, and discrepancies in spatial and temporal resolution between macro-scale thermal fields and micro-scale grain evolution.

Consequently, rigorous and multi-level validation remains indispensable. This should include DOE-based testing, robustness evaluations under boundary and near-boundary conditions, and consistency checks across transient process cycles. Together, these validation steps help ensure that the resulting predictive framework remains accurate, stable, and reliable during deployment in advanced digital manufacturing environments.

Supplementary Materials: The following supporting script can be downloaded at <https://gitlab.com/ait-lkr/directchillcasting>; CFD open-source code for numerical simulations (accessed on 15 February 2026).

Author Contributions: S.M.: methodology, investigation, scripting (CVAE), data curation, validation, writing—original draft preparation, proofreading, editing, visualization. A.M.H.: conceptualization, methodology, writing—original draft preparation, software (data models), validation,

writing—review, editing, visualization. S.E.: supervision, writing—original draft preparation, proofreading, editing. M.H.: investigation, data curation, validation, writing—original draft preparation, proofreading. R.G.V.: investigation, CFD simulations, validation, writing—original draft preparation, proofreading, visualization. All authors have read and agreed to the published version of the manuscript.

Funding: This research is financially supported by the Austrian Institute of Technology (AIT) under the UF2025 funding program, and by the Austrian Research Promotion Agency (FFG) through the opt1mus project (FFG No. 899054).

Data Availability Statement: The sharing of raw data required to reproduce the case studies are considered upon request by readers.

Acknowledgments: The authors gratefully acknowledge the technical and financial support provided by the Austrian Federal Ministry for Innovation, Mobility and Infrastructure, the Federal State of Upper Austria, and the Austrian Institute of Technology (AIT). Special thanks are extended to David Blacher and Johannes Kronsteiner for their valuable contributions to this research.

Conflicts of Interest: Authors Sofija Milicic, Amir M. Horr, Manuel Hofbauer, and Rodrigo Gómez Vázquez are employed by the company LKR Light Metals Technologies Ranshofen, Austria. Author Stefanie Elgeti is employed by Institute of Lightweight Design and Structural Biomechanics, TU Wien, Chair for Computational Analysis of Technical Systems, RWTH Aachen University, Germany, and DCSE, TU Delft, The Netherlands.

References

1. Ciccone, F.; Bacciaglia, A.; Ceruti, A. Optimization with artificial intelligence in additive manufacturing: A systematic review. *J. Braz. Soc. Mech. Sci. Eng.* **2023**, *45*, 303. [\[CrossRef\]](#)
2. Chen, K.; Zhang, P.; Yan, H.; Chen, G.; Sun, T.; Lu, Q.; Chen, Y.; Shi, H. A review of machine learning in additive manufacturing: Design and process. *Int. J. Adv. Manuf. Technol.* **2024**, *135*, 1051–1087. [\[CrossRef\]](#)
3. Chia, H.Y.; Wu, J.; Wang, X.; Yan, W. Process parameter optimization of metal AM: Review & outlook. *J. Mater. Inf.* **2022**, *2*, 16. [\[CrossRef\]](#)
4. Menold, T.; Onuseit, V.; Buser, M.; Haas, M.; Bär, N.; Michalowski, A. Laser material processing optimization using Bayesian optimization. *Light Adv. Manuf.* **2024**, *5*, 32. [\[CrossRef\]](#)
5. Liu, X.; Han, X.; Wang, H.; Liu, G. A modelling and updating approach of digital twin based on surrogate model to rapidly evaluate product performance. *Int. J. Adv. Manuf. Technol.* **2023**, *129*, 5059–5074. [\[CrossRef\]](#)
6. Tang, Y.; Sajadi, P.; Rahmani Dehaghani, M.; Wang, G.G. A systematic online update method for reduced-order-model-based digital twin. *J. Intell. Manuf.* **2025**, *36*, 5495–5523. [\[CrossRef\]](#)
7. Horr, A.M.; Drexler, H. Real-Time Models for Manufacturing Processes: How to Build Predictive Reduced Models. *Processes* **2025**, *13*, 252. [\[CrossRef\]](#)
8. Mallor, C.; Lani, S.; Zambrano, V.; Ghasemi-Tabasi, H.; Calvo, S.; Burn, A. Integrating 3D printing, simulations and surrogate modelling: A comprehensive study on additive manufacturing focusing on a metal twin-cantilever benchmark. *Adv. Ind. Manuf. Eng.* **2025**, *10*, 100162. [\[CrossRef\]](#)
9. Balu, A.; Sarkar, S.; Ganapathysubramanian, B.; Krishnamurthy, A. Physics-aware machine learning surrogates for real-time manufacturing digital twin. *Manuf. Lett.* **2022**, *34*, 71–74. [\[CrossRef\]](#)
10. Leng, J.; Zuo, K.; Xu, C.; Zhou, X.; Zheng, S.; Kang, J.; Liu, Q.; Chen, X.; Shen, W.; Wang, L.; et al. Physics-informed machine learning in intelligent manufacturing: A review. *J. Intell. Manuf.* **2025**, *37*, 2215–2257. [\[CrossRef\]](#)
11. Pan, L.; Li, G.; Zhu, T.; Liu, D.; Wang, Y.; Lu, Y. Physics-Informed Machine Learning in Design and Manufacturing: Status and Challenges. *ASME J. Comput. Inf. Sci. Eng.* **2025**, *25*, 120804. [\[CrossRef\]](#)
12. Urich, B.; Pfeifer, N.; Schäfer, M.; Theile, O.; Rahm, E. Physics-informed deep learning to quantify anomalies for real-time fault mitigation in 3D printing. *Appl. Intell.* **2024**, *54*, 4736–4755. [\[CrossRef\]](#)
13. Yang, S.; Peng, S.; Guo, J.; Wang, F. A review on physics-informed machine learning for monitoring metal additive manufacturing process. *Adv. Manuf.* **2024**, *1*, 8. [\[CrossRef\]](#)
14. Anonyuo, S.; Kwakye, J.M.; Ozowe, W. A Review of Quality Control and Process Optimization in High-Volume Semiconductor Manufacturing. *World J. Eng. Technol. Res.* **2024**, *3*, 22–27. [\[CrossRef\]](#)
15. Lu, Z.; Ren, N.; Xu, X.; Li, J.; Panwisawas, C.; Xia, M.; Dong, H.; Tsang, E.; Li, J. Real-time prediction and adaptive adjustment of continuous casting based on deep learning. *Commun. Eng.* **2023**, *2*, 34. [\[CrossRef\]](#)

16. Knoth, S. Statistical Process Control. In *Applied Quantitative Finance*; Springer: Berlin/Heidelberg, Germany, 2002. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Montgomery, D.C. The use of statistical process control and design of experiments in product and process improvement. *IIE Trans.* **1992**, *24*, 4–17. [\[CrossRef\]](#)
18. Haridy, S.; Gouda, S.A.; Wu, Z. An integrated framework of statistical process control and design of experiments for op-timizing wire electrochemical turning process. *Int. J. Adv. Manuf. Technol.* **2011**, *53*, 191–207. [\[CrossRef\]](#)
19. Dogan, O.; Areta Hiziroglu, O. Empowering Manufacturing Environments with Process Mining-Based Statistical Process Control. *Machines* **2024**, *12*, 411. [\[CrossRef\]](#)
20. Montgomery, D.C. The 100th anniversary of the control chart. *J. Qual. Technol.* **2024**, *56*, 2–4. [\[CrossRef\]](#)
21. Best, M.; Neuhauser, D. Walter A Shewhart, 1924, and the Hawthorne factory. *BMJ Qual. Saf.* **2006**, *15*, 142–143. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Bradford, P.G.; Miranti, P.J. Information in an Industrial Culture: Walter A. Shewhart and the Evolution of the Control Chart, 1917–1954. *Inf. Cult.* **2019**, *54*, 179–219. [\[CrossRef\]](#)
23. Luceño, A. Control Charts. In *International Encyclopedia of Statistical Science*; Lovric, M., Ed.; Springer: Berlin/Heidelberg, Germany, 2025. [\[CrossRef\]](#)
24. Swamidass, P.M. SHEWHARTCONTROL CHARTS. In *Encyclopedia of Production and Manufacturing Management*; Swamidass, P.M., Ed.; Springer: New York, NY, USA, 2000. [\[CrossRef\]](#)
25. Radanliev, P. Artificial intelligence: Reflecting on the past and looking towards the next paradigm shift. *J. Exp. Theor. Artif. Intell.* **2025**, *37*, 1045–1062. [\[CrossRef\]](#)
26. Koren, Y.; Heisel, U.; Jovane, F.; Moriwaki, T.; Pritschow, G.; Ulsoy, G.; Van Brussel, H. Reconfigurable Manufacturing Systems. *CIRP Ann.* **1999**, *48*, 527–540. [\[CrossRef\]](#)
27. Coons, S.A. An outline of the requirements for a computer-aided design system. In *Proceedings of the May 21–23, 1963, Spring Joint Computer Conference (AFIPS '63 (Spring))*; Association for Computing Machinery: New York, NY, USA, 1963; pp. 299–304. [\[CrossRef\]](#)
28. Marr, D.; Hildreth, E. Theory of edge detection. *Proc. Biol. Sci.* **1980**, *207*, 187–217. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Buchanan, B.G.; Shortliffe, E.H. *Rule Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project*; Addison-Wesley Series in Artificial Intelligence; Addison-Wesley Longman Publishing: Boston, MA, USA, 1984.
30. Zadeh, L.A. The concept of a linguistic variable and its application to approximate reasoning-I. *Inf. Sci.* **1975**, *8*, 199–249. [\[CrossRef\]](#)
31. Kanet, J.J.; Adelsberger, H.H. Expert systems in production scheduling. *Eur. J. Oper. Res.* **1987**, *29*, 51–59. [\[CrossRef\]](#)
32. Slam, M.T.; Sepanloo, K.; Woo, S.; Woo, S.H.; Son, Y.-J. A Review of the Industry 4.0 to 5.0 Transition: Exploring the Inter-section, Challenges, and Opportunities of Technology and Human–Machine Collaboration. *Machines* **2025**, *13*, 267. [\[CrossRef\]](#)
33. Oks, S.J.; Jalowski, M.; Lechner, M.; Mirschberger, S.; Merklein, M.; Vogel-Heuser, B.; Möslin, K.M. Cyber-Physical Systems in the Context of Industry 4.0: A Review, Categorization and Outlook. *Inf. Syst. Front* **2024**, *26*, 1731–1772. [\[CrossRef\]](#)
34. van-Lopik, K.; Hayward, S.; Grant, R.; McGirr, L.; Goodall, P.; Jin, Y.; Price, M.; West, A.A.; Conway, P.P. A review of design frameworks for human-cyber-physical systems moving from industry 4 to 5. *IET Cyber-Phys. Syst. Theory Appl.* **2023**, *9*, 169–183. [\[CrossRef\]](#)
35. Krugh, M.; Mears, L. A complementary cyber-human systems framework for industry 4.0 cyber-physical systems. *Manuf. Lett.* **2018**, *15*, 89–92. [\[CrossRef\]](#)
36. Van-Lopik, K. Considerations for the Design of Next-Generation Interfaces to Support Human Workers in Industry 4.0. Doctoral Thesis, Loughborough University, Loughborough, UK, 2019. [\[CrossRef\]](#)
37. Nakagawa, E.Y.; Antonino, P.O.; Schnicke, F.; Capilla, R.; Kuhn, T.; Liggesmeyer, P. Industry 4.0 reference architectures: State of the art and future trends. *Comput. Ind. Eng.* **2021**, *156*, 107241. [\[CrossRef\]](#)
38. Sharpe, R.; van Lopik, K.; Neal, A.; Goodall, P.; Conway, P.P.; West, A.A. An industrial evaluation of an Industry 4.0 reference architecture demonstrating the need for the inclusion of security and human components. *Comput. Ind.* **2019**, *108*, 37–44. [\[CrossRef\]](#)
39. Nguyen, H.; Ganix, N.; Ion, L. Human-centred design in industry 4.0: Case study review and opportunities for future research. *J. Intell. Manuf.* **2022**, *33*, 35–76. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Manzoni, A.; Quarteroni, A.; Rozza, G. Computational Reduction for Parametrized PDEs: Strategies and Applications. *Milan J. Math.* **2012**, *80*, 283–309. [\[CrossRef\]](#)
41. Dissanayake, M.W.M.G.; Phan-Thien, N. Neural-network-based approximations for solving partial differential equations. *Commun. Numer. Methods Eng.* **1994**, *10*, 195–201. [\[CrossRef\]](#)
42. Lagaris, I.E.; Likas, A.; Fotiadis, D.I. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Trans. Neural Netw.* **1998**, *9*, 987–1000. [\[CrossRef\]](#) [\[PubMed\]](#)

43. Psychogios, D.C.; Ungar, L.H. A hybrid neural network-first principles approach to process modeling. *AIChE J.* **1992**, *38*, 1499–1511. [\[CrossRef\]](#)
44. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Comput. Phys.* **2019**, *378*, 686–707. [\[CrossRef\]](#)
45. Rouhollahi, A.; Rismanian, M.; Ebrahimi, A.; Ilegbusi, O.J.; Nezami, F.R. Prediction of directional solidification in freeze casting of biomaterial scaffolds using physics-informed neural networks. *Biomed. Phys. Eng. Express* **2024**, *10*, 6. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Würth, T.; Krauß, C.; Zimmerling, C.; Kärger, L. Physics-informed neural networks for data-free surrogate modelling and engineering optimization—An example from composite manufacturing. *Mater. Des.* **2023**, *231*, 112034. [\[CrossRef\]](#)
47. Song, R.; Eo, S.; Lee, M.; Lee, J. Physics-informed machine learning for optimizing the coating conditions of blade coating. *Phys. Fluids* **2022**, *34*, 082112. [\[CrossRef\]](#)
48. Horr, A.M. Notes on New Physical & Hybrid Modelling Trends for Material Process Simulations. *J. Phys. Conf. Ser.* **2020**, *1603*, 012008. [\[CrossRef\]](#)
49. Horr, A.; Blacher, D.; Vázquez, G. On Performance of Data Models and Machine Learning Routines for Simulations of Casting Processes. *BHM Berg Hüttenmänn. Monatshefte* **2025**, *170*, 28–36. [\[CrossRef\]](#)
50. Horr, A.M. Real-Time Modeling for Design and Control of Material Additive Manufacturing Processes. *Metals* **2024**, *14*, 1273. [\[CrossRef\]](#)
51. Horr, A.M.; Hartmann, M.; Haunreiter, F. Innovative Data Models: Transforming Material Process Design and Optimization. *Metals* **2025**, *15*, 873. [\[CrossRef\]](#)
52. Horr, A.M. Real-time Modelling and ML Data Training for Digital Twinning of Additive Manufacturing Processes. *BHM Berg Hüttenmänn. Monatshefte* **2024**, *169*, 48–56. [\[CrossRef\]](#)
53. Kingma, D.P.; Welling, M. An Introduction to Variational Autoencoders. *Found. Trends Mach. Learn.* **2019**, *12*, 307–392. [\[CrossRef\]](#)
54. Doersch, C. Tutorial on Variational Autoencoders. *arXiv* **2016**, arXiv:1606.05908. [\[CrossRef\]](#)
55. Ivanov, O.; Figurnov, M.; Vetrov, D. Variational Autoencoder with Arbitrary Conditioning. *arXiv* **2018**, arXiv:1806.02382. [\[CrossRef\]](#)
56. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv* **2015**, arXiv:1505.04597. [\[CrossRef\]](#)
57. Bennon, W.D.; Incropera, F.P. A continuum model for momentum, heat and species transport in binary solid-liquid phase change systems—I. Model formulation. *Int. J. Heat Mass Transf.* **1987**, *30*, 2161–2170. [\[CrossRef\]](#)
58. Lebon, B. directChillFoam: An OpenFOAM application for direct-chill casting. *J. Open Source Softw.* **2023**, *8*, 4871. [\[CrossRef\]](#)
59. Thévoz, P.; Desbiolles, J.L.; Rappaz, M. Modelling of equiaxed microstructure formation in casting. *Metall. Mater. Trans. A* **1989**, *20*, 311–322. [\[CrossRef\]](#)

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.