Article

Digital Twin-Driven Multi-Factor Production Capacity Prediction for Discrete Manufacturing Workshop

Hu Cai [1], Jiafu Wan [1,*] and Baotong Chen [2,3]

1   Guangdong Provincial Key Laboratory of Precision Equipment and Manufacturing Technology, South China University of Technology, Guangzhou 510641, China; 201910100263@mail.scut.edu.cn
2   Key Laboratory of Metallurgical Equipment and Control Technology, Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China; mecheney@wust.edu.cn
3   Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China
*   Correspondence: mejwan@scut.edu.cn

Abstract: Traditional capacity forecasting algorithms lack effective data interaction, leading to a disconnection between the actual plan and production. This paper discusses the multi-factor model based on a discrete manufacturing workshop and proposes a digital twin-driven discrete manufacturing workshop capacity prediction method. Firstly, this paper gives a system framework for production capacity prediction in discrete manufacturing workshops based on digital twins. Then, a mathematical model is described for discrete manufacturing workshop production capacity under multiple disturbance factors. Furthermore, an innovative production capacity prediction method, using the "digital twin + Long-Short-Term Memory Network (LSTM) algorithm", is presented. Finally, a discrete manufacturing workshop twin platform is deployed using a commemorative disk custom production line as the prototype platform. The verification shows that the proposed method can achieve a prediction accuracy rate of 91.8% for production line capacity. By integrating the optimization feedback function of the digital twin system into the production process control, this paper enables an accurate perception of the current state and future changes in the production system, effectively evaluating the production capacity and delivery date of discrete manufacturing workshops.

Keywords: digital twins; industrial big data; multi-factor model; discrete manufacturing workshop; productivity prediction

1. Introduction

The rapid development of information and communication technology has significantly facilitated the digital transformation of enterprises. This process has also promoted the rapid development of enabling technologies, such as the industrial Internet, digital twins, and industrial big data. Moreover, the increasing demand for personalized products will affect the system, production lines, and products. The enterprise digital transformation is developing towards multi-variety, small-batch, and personalized production, which promotes the realization of intelligent manufacturing, product life cycle monitoring, and "Made in China 2025" [1]. Aiming at multi-level spans of time/space scales of components in the production line cyber–physical system, order-driven and random disturbances interact in the production process. Moreover, the physical resource allocation/execution and information decision-making reconstruction are loosely related to the control link [2].

The production mode in discrete manufacturing workshops is different from process production. The production process is usually decomposed into multiple processing tasks to complete, and each of them is applied to specific equipment or resources within the enterprise. It should be noted that tasks with the same process are assigned to one device for processing. At the same time, the production process of discrete manufacturing workshops needs to adapt to changes in customer needs, product forms, and production factors. In

discrete workshops, many factors affect production planning, including changes in product types, production delivery dates, and equipment status, thereby adding complexity to the formulation of production plans and the prediction of actual factory capacity. In the past production, factory scheduling and delivery forecasts were mainly evaluated by scheduling personnel based on experience and historical data. However, with the ever-changing demand for product production, the management of production factors, production activity planning, and production process control are becoming increasingly complex. Managing, controlling, predicting, analyzing, and optimizing efficiently, reasonably, and in real time are among the important issues facing discrete manufacturing workshops. Furthermore, before a company builds a new production base and processes new products, realizing the company's production capacity forecast, delivery date evaluation, and rational production arrangement must be the focus of the company's work.

Discrete manufacturing workshop production activity planning and production capacity evaluation typically involve two stages [3,4]. In the first stage, experienced engineers manage personnel, materials, equipment, processes, and other information in the discrete manufacturing workshop. Material storage, planning arrangements, capacity forecasting, etc., are all evaluated based on personnel experience and historical information and stored on paper media. Information statistics, transmission, query, and analysis are difficult, leading to a lower accuracy rate. However, with the introduction of Enterprise Resource Planning (ERP), Manufacturing Execution System (MES) systems, and computer technology, data statistics, archival, and transmission capabilities have been significantly improved. Since information processing and analysis capabilities are continuously improved, factory informatization has entered the second stage [5]. According to the material requirement plan, process route, and man-hour quota, the time and workload required for production capacity are given and compared with the capacity provided by the manufacturing workshop. This evaluation shows whether completing the material requirement plan matches the ability. If the utilization rate of the equipment is unacceptable, the master production plan should be adjusted, and the material requirements plan should be regenerated. During this stage, the system can realize production and processing resource constraints, production process allocation, production equipment monitoring, etc. However, due to the lack of real-time interaction and information fusion capabilities, production capacity prediction and delivery dates rely more on historical data for model evaluation [6].

Before building new factories and processing new products, problems such as a lack of historical data are exposed, making it impossible to predict workshop production capacity accurately. Combining the current requirements of intelligent manufacturing systems, the introduction of digital workshops signifies a change in the elements of traditional manufacturing systems. Moreover, the concept of digital twin workshops has also emerged as the times require [7]. In the digital twin manufacturing workshop, the data of production factors such as personnel, equipment, materials, and environment in the physical workshop are fully connected to the information world through the workshop Internet of Things (IoT), RFID, and other information means, realizing mutual interconnection and data sharing. Driven by comprehensive production factor data, the simulation, evaluation, and analysis functions of the virtual workshop and workshop service system can optimize the behavior of each factor while considering the status of other factors. This approach supports the linkage and optimal combination of factors necessary to ensure smooth production operations [8]. The data in the physical space mainly refer to the real data generated by the physical entities of the workshop. Conversely, the data in the information space mainly refer to the data related to the virtual workshop and the workshop service system. These data are not directly collected from the physical space but from virtual workshop data which is based on digital twin data model, algorithm deduction, and system derivation [9]. With the introduction of digital twin technology, a multi-physical, multi-scale, and multi-probability twin of a discrete manufacturing workshop is established through digital means to achieve deep integration and real-time interaction between physical space and information space, offering a basis for data-driven workshop capacity prediction.

Considering the characteristics of discrete manufacturing workshops, such as the production of many types of products, small batches, and frequent production changes, this paper discusses the problems of inaccurate production capacity forecasts and unfulfilled delivery dates during the factory production process. Based on predictive data, it is possible to optimize production scheduling methods in digital twin systems and provide follow-up suggestions. The modeling method of discrete manufacturing workshops driven by twin data and the establishment process of factor conditions are considered based on the problems faced by discrete manufacturing workshops. In addition, the neural network algorithm to realize accurate predictions of factory production capacity and delivery management is studied based on the interactive fusion of twin space and twin data. Finally, the feasibility of this method is experimentally verified. The main contributions of this paper are in the following three aspects:

1. Discussing the current production capacity forecasting methods and problems faced by discrete manufacturing workshops and proposing a digital twin-driven discrete manufacturing workshop capacity forecasting system framework;
2. Innovating the modeling process of discrete manufacturing workshops driven by twin data and proposing a time series prediction algorithm for workshop production capacity based on Long Short-Term Memory (LSTM);
3. Taking the lead in building a digital twin platform for capacity prediction in discrete manufacturing workshops and conducting digital deployment in actual manufacturing workshops to verify the accuracy of the prediction model.

The subsequent chapters of this paper are arranged as follows: The related work is described in the second section. Section 3 considers the production mode, management mode, and current production capacity forecasting method of the discrete manufacturing workshop. Moreover, the system framework for capacity forecasting in the discrete manufacturing workshop is presented. In Section 4, this paper establishes the capacity problem model for discrete workshops driven by digital twins. Section 5 presents the design of production capacity prediction algorithms driven by twin data, along with a discussion of the constraints and solutions involved in the process of building the information space of discrete manufacturing workshops using digital twin technology. Section 6 conducts an experiment to validate the feasibility of the proposed approach. Finally, this paper concludes with a summary of the key findings and contributions.

2. Related Works

This section introduces the current related research on aspects of production capacity prediction in discrete workshops, workshop production capacity control driven by twin data, and workshop production capacity prediction algorithms.

2.1. Capacity Forecasting for Discrete Workshops

In a discrete manufacturing workshop, the production process of products is typically divided into multiple processing tasks. Each task is applied to specific equipment or resources within the enterprise, ensuring that tasks with the same process are assigned to the same equipment for processing. The production process route and equipment usage are flexible and variable, and there are many changes in product design, processing requirements, order quantity, production lead time, and other aspects during the production process. The dynamic changes in product types and equipment integration in discrete manufacturing workshops involve relatively complex product processes and equipment processing state constraints, requiring multiple processing workstations for processing and component assembly; the product shape and structure are highly dependent on customer preferences, leading to a high degree of customization.

Ozturk et al. [10] proposed a branch-and-bound algorithm designed specifically to solve parallel-batch scheduling problems with different processing times, release times, and unit sizes, ensuring that the number of job batches does not exceed the equipment capacity and that the processing time is determined based on the maximum processing time within

the batch. Zhou et al. [11] introduced a knowledge graph-based framework for discrete manufacturing workshop resource allocation, integrating workshop implicit engineering knowledge to support the optimal resource allocation method. During the production and operation process of enterprises, many different varieties and series of products are produced simultaneously, resulting in complex and diverse products and frequent production changes. From the perspective of product processing, the production process in a discrete manufacturing workshop is a complex process formed by the parallel or serial connection of various component-processing subprocesses. This process contains complex constraints and disturbance factors. Therefore, the process control of discrete manufacturing enterprises becomes more complex and variable. Table 1 provides an overview of the research methods employed by different scholars and highlights their advantages.

In discrete manufacturing workshops, many factors affect production planning, resulting in a very complex formulation of production plans. Workshop capacity prediction and production lead time are often estimated by the planning and scheduling person based on experience. However, any changes in the production process plan frequently lead to insufficient capacity estimation and delayed delivery times. To meet these needs, enterprises must be equipped with highly customized supply networks and create resilient corporate risk management capabilities to improve the decision support system for the design, scheduling, control, and collaborative management of digital manufacturing [12].

Dengiz et al. [13] pointed out that simulation modeling is one of the most useful techniques for analyzing and evaluating the dynamic behavior of complex manufacturing systems. They employed virtual simulation analysis modeling to improve productivity in the automotive industry. Benotsmane et al. [14] used numerical and graphical modeling and simulation, 3D designs, and Tabu search in trajectory space to evaluate and enhance productivity through collaborative robot simulation and artificial intelligence methods. A. Florescu and S. A. Barabas [15] utilized simulation in the control system of intelligent manufacturing systems to analyze current parameters and predict the requirements of future processing scenarios. In addition, in the customized needs of customers, the complexity of the process and high-quality requirements necessitate the evaluation of workshop production capacity based on multiple factors, such as specific product processes, production equipment status, and disturbance factors. Relying on on-site operators for estimation may lead to significant deviations in processing time and production capacity estimation due to insufficient personnel experience and on-site consideration. Consequently, the goal of achieving timely delivery may fail. S. Kim and K. Ryu [16] proposed the definition of AI-based statistical analysis technology, which utilizes on-site information to estimate the maximum workshop capacity for the proposed process and improve the competitiveness of the mold industry. Wang et al. [17] proposed a proactive manufacturing resource allocation (PMRA) method based on intelligent factory production performance prediction. This approach avoids production interruptions or performance degradation issues based on real-time analysis and an accurate prediction of production performance.

Table 1. Research on production capacity forecasting.

Author	Characteristic	Advantage
Ozturk et al. [10]	Production capacity forecast based on equipment capacity, processing time, etc.	For a certain production line, it can quickly predict the production capacity.
Wang et al. [17]	Accurate predictions based on real-time analysis and production performance.	Real-time dynamic analysis based on equipment and production line status.
Zhou et al. [11]	A resource allocation framework for discrete manufacturing workshops based on knowledge graphs for capacity prediction.	Coordination of all resources in the workshop and evolution of the current production capacity forecast of the workshop based on the knowledge graph.

Table 1. *Cont.*

Author	Characteristic	Advantage
Dengiz et al. [13]	Improving productivity in the automotive industry using hybrid simulation–analytical modeling.	Analysis of the bottleneck of automobile production capacity based on the digital twin system, optimization of the design, and increase in production capacity.
This article	Establishing a prediction model for "digital twin data + LSTM network". Iteratively optimizing the system through twin data and systems and dynamically fitting the production capacity forecast under different states. It has better applicability and reliability.	Full use of the advantages of a virtual and real fusion of digital twin systems and real production lines. With the support of digital twin data, consideration of the constraints of the workshop, and the analysis of production performance under different states and resources. Combined with the LSTM algorithm, facilitation of real-time and accurate analysis of production line capacity.

Hence, for discrete manufacturing enterprises, effectively estimating production capacity and reasonably predicting delivery time based on production equipment elements, process requirements, and constraints is a critical issue before constructing discrete workshops and undertaking product processing. Once production anomalies occur, production adjustment demands are often released, and production decisions are often based on historical production information, which may result in production interruptions or performance degradation. Capacity prediction is a complex time series prediction problem that considers multiple factor disturbances.

2.2. Workshop Capacity Control Driven by Digital Twins

Discrete manufacturing workshops encounter many challenges and uncontrollable factors in production factor management, production activity planning management, and production process control. When building new factories and processing new products, achieving accurate production control processes is difficult due to a lack of historical data as a reference. Digital twinning has brought digital engineering modeling and simulation into a new era [18]. Serving as the digital representation of real-world entities and systems, the information space within digital twins can realize the multi-dimensional, multi-scale, and multi-probability simulation process of physical entities. Through the interaction and fusion of twin data, digital twins facilitate the understanding, prediction, optimization, and control of real entities or systems [19,20].

The connotations, reference models, applications, and research issues of digital twin-driven intelligent manufacturing have been proposed in [21]. This application focuses on continuous optimization, proactive maintenance, and continuous data processing throughout the production process. Vachálek et al. [22] highlighted that the digital twin of industrial production lines could support the existing production structure within the automotive industry and utilize resources effectively through expanded production and planning strategies. Ma et al. [23] proposed a digital twin-driven production management system for production workshops, enabling the dynamic simulation and optimization of the production process of the manufacturing industry while achieving real-time synchronization, high fidelity, and real virtual integration in network physical production. In response to constantly changing customer demands, rising resource costs, and growing uncertainty, Kunath et al. [24] integrated the digital twin of manufacturing systems into decision support systems to enhance the order management process. In summary, digital twins achieve statistical, analytical, and optimized decision-making based on twin data through real-time interaction and fusion between the physical and information worlds. By utilizing continuous simulation and iterative input data and constraint conditions, digital twins offer an effective solution to the problem of inaccurate production capacity prediction prior to factory construction and product production and processing.

Due to the asynchronous data between the physical and information spaces in the workshop, a significant delay exists in production process control, resulting in a disconnection from actual production. Defects will be observed when using basic process data (key elements of Industry 4.0) for the fully automated data collection process, as well as the evaluation, quantification, and analysis of collected data. According to the superposition of digital twins and new-generation information technology, the discrete manufacturing physical and digital workshops have achieved bidirectional real mapping and real-time interaction capabilities, fully realizing the management mode encompassing all elements, processes, and businesses within the entire discrete manufacturing workshop. Driven by twin data, continuous iterative optimization is achieved by setting product process routes, production line constraints, production factor conditions, and production process control to achieve the optimal operation mode of workshop production and control. The virtual workshop model and related information can be overlaid and interacted with the physical workshop in real time, thereby simulating the operation process of the physical workshop based on historical time series data. Digital chain technology enables seamless integration, real-time interaction, and fusion between the virtual workshop and the physical workshop, which is expected to drive the production capacity evaluation and prediction of the workshop.

2.3. Prediction Algorithm for Production Capacity in Discrete Workshops

To address the problem of workshop capacity fluctuations with attributes such as large scale, strong NP-hard, non-convex optimization, and randomness, a combined approach employing a digital twin model and a digital twin data-driven approach is proposed to study production capacity prediction algorithms. These algorithms exhibit time sensitivity to meet fast and effective solution requirements, which is particularly crucial. In the past production, since yield prediction is a time series data analysis problem, the main methods included linear regression, exponential smoothing, and time series smoothing. However, with the increasing accuracy requirements of production capacity prediction models, advanced algorithms such as heuristic graph search, simulated annealing, genetic algorithms, and neural networks have emerged. Among them, in the discrete workshop, the ARIMA algorithm is suitable for predicting and analyzing the production capacity of a single product during stable equipment operation, maintaining data stability without production state iteration; RNN can effectively iterate process data and is suitable for processes where the state of products and equipment in discrete workshops is constantly changing. However, due to gradient explosion, convergence cannot be achieved, which is shown in Table 2.

Table 2. Comparison of production capacity prediction algorithms.

| Algorithm Name	Characteristic	Advantage	Disadvantage
ARIMA [25]	Quickly captures time series data features	Strong pertinence, good robustness	Need for smoothing the data
Holt–Winters [26]	Effectively predicts data with seasonal trends	Automatically shields the influence of abnormal data; the process is simple	False positives for periodic anomaly data
GBM-class regression algorithm [27]	Overfitting can be controlled by controlling the number of iterations	Model learning with a high degree of freedom and high precision	Promotion and application complexity are high
RNN [28]	Handles data input of arbitrary length	High learning freedom and good customization	Long-term dependence on data, resulting in gradient disappearance
LSTM [29]	Solves the long-term dependency problem of RNNs	With long-term memory function	High data volume requirements, limited optimization methods

2.4. Summary of Relevant Research

The above research mainly elaborates on discrete workshop capacity prediction, digital twin modeling, driving decision-making, and capacity prediction algorithms. This clearly illustrates the value and prospects of digital twins in workshop control. It also elaborates on the types and adaptability of production prediction algorithms. However, there is a lack of correlation and architectural systems for establishing digital twin models and capacity analysis. No effective algorithm has been developed to demonstrate the value of data analysis and their application in discrete workshop capacity prediction in digital twins.

This paper establishes a digital twin workshop to simulate the operation and evolution of the workshop in different stages, such as discrete manufacturing production factor management, production plan management, and production process control driven by twin data. Then, it achieves applications, including evaluation, optimization, prediction, and traceability. By integrating production process control with the optimization feedback functions of the twin system, it becomes possible to accurately perceive the current state and future changes in the system, thereby effectively evaluating the workshop's production capacity and delivery time.

3. System Framework for Capacity Prediction in Discrete Manufacturing Workshop

Discrete manufacturing workshops often allocate and schedule production based on the specific process and equipment status of products, resulting in many constraints on production scheduling and abnormally complex production process control. During the production process, there are many non-standard products due to the large variety of product types, and most production enterprises often have on-site workers for flexible adaptation [30]. Consequently, data such as production evaluation, scheduling, production capacity prediction, and delivery time prediction rely on the abilities and experience of employees, leading to inaccuracies in predicting the actual production capacity of the workshop. In the production process of discrete manufacturing workshops, a variety of heterogeneous data sources, such as customer data, material data, production activity planning data, equipment status, and time-varying production process control, are involved. Currently, data analysis and processing are only performed manually or through ERP, MES, and other application systems. However, these approaches lack real-time interaction and information fusion, thereby limiting effective control over current production factors, production activity plans, and production process control within the workshop.

The digital twin of discrete manufacturing workshops represents a digital mapping of discrete workshops, as shown in Figure 1, providing a system framework for predicting workshop capacity driven by digital twins. Based on the digital twin five-dimensional model theory, the discrete manufacturing workshop twin system involves physical workshops, virtual workshops, service systems, twin data, and connected services. The construction process of the digital twin system includes the construction of production factors, such as the human–machine material environment in the workshop, a production activity plan and process construction that organizes and distributes production factors reasonably, and system modeling processes, i.e., production system monitoring, analysis, and optimization. The digital twin system depicts all elements, processes, and business of the physical workshop, along with real-time interaction and system integration based on twin data. Through a full-factor modeling process based on the twin system, precise characterization of production factors in discrete manufacturing workshops can be achieved, leading to the establishment of a twin model that exhibits virtual and real consistencies. Additionally, a dynamic modeling process based on models and twin data can effectively solve the problem of twin data-driven production process state identification and dynamic iterative modeling. This process enables equipment fault diagnosis and predictive maintenance through iterative analysis based on the real-time status of the production line. Furthermore, the simulation analysis and decision-making of the production process are achieved based on the twin model and data fusion-driven modeling process by establishing constraints and coupling relationships among production factors, models, and data.
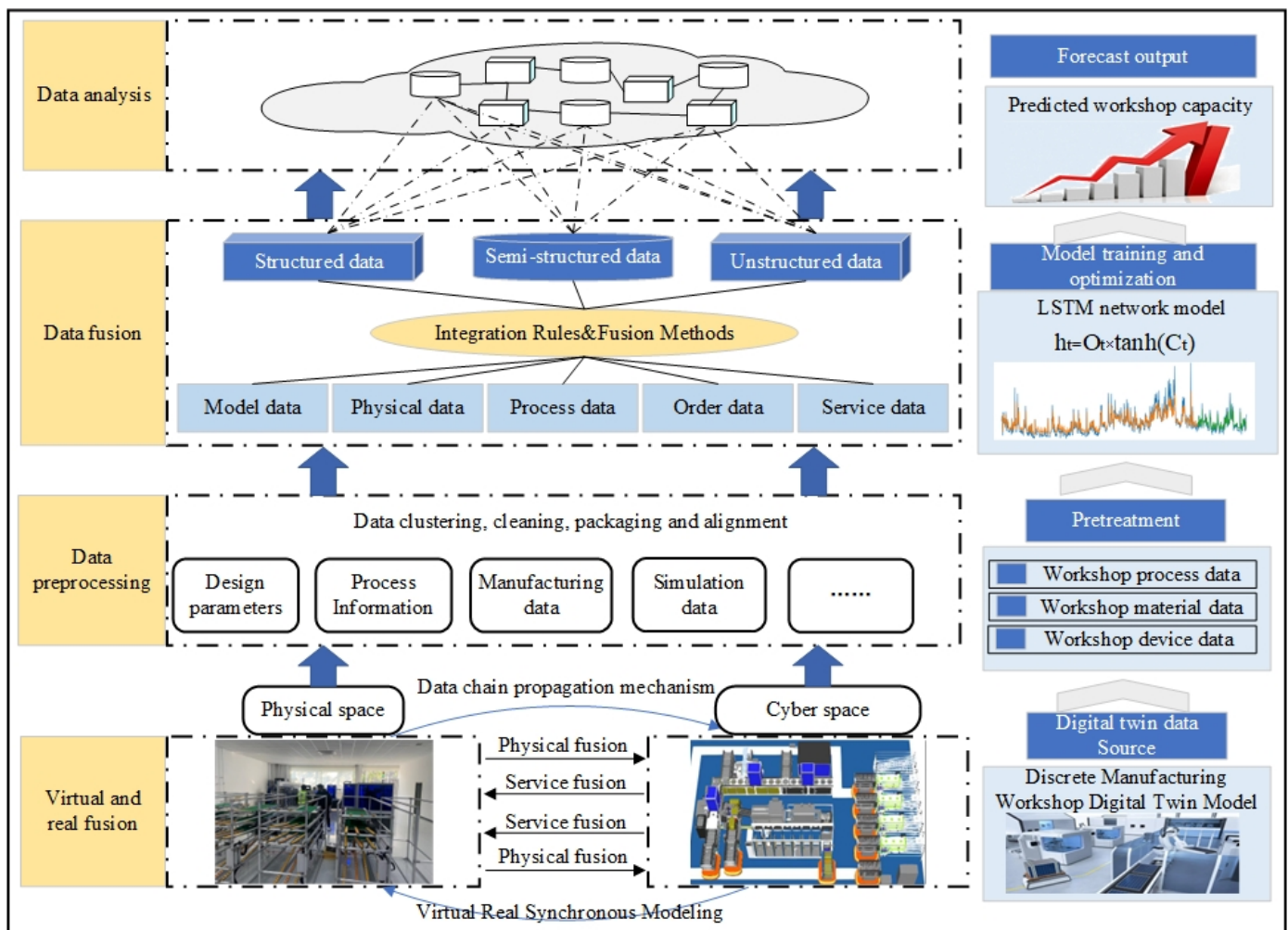
Figure 1. Workshop capacity prediction framework driven by digital twin data.

The precise prediction of production capacity for discrete manufacturing workshops is shown in Figure 1. Firstly, a digital twin of discrete workshops is constructed using the software of Process Simulate 16.0.0. This architecture enables real-time interaction between information space and physical space through the utilization of digital chain technology. At this level, virtual and physical space production lines achieve physical fusion based on the same 3D and behavioral models. Based on mutual control between virtual space and physical space, bidirectional data transmission and service integration are achieved. Secondly, the data related to workshop production (cluster analysis, cleaning, and packaging) are preprocessed, which mainly includes workshop process data, workshop material data, and workshop equipment status. Then, the data are fused, and a prediction model is established using a combination of "digital twin data + LSTM network". Finally, an accurate prediction of workshop production capacity is achieved through a series of data analysis techniques. By integrating all elements, processes, and business control in the production process, as well as the optimization feedback function of the twin system, it is possible to accurately perceive the current state and future changes in the system.

4. Discrete Shop Capacity Model Driven by Digital Twins

The discrete workshop capacity analysis based on digital twins is built on a data-driven twin system simulation model. The mathematical relationship between products and production capacity can be established by simulating the production process of twin systems and iterating the data. The first step in this analysis is to construct a model for discrete workshop production capacity.

The occurrence of production capacity disturbance factors in discrete manufacturing workshops is random and variable. Twin data can be used in the model based on the fast simulation iteration ability of twin models to iterate production line operation information, collect and analyze production process data, and analyze data based on material pulling. Neural networks and other methods can be used to model the impact of disturbance factors on the production status and then achieve the model establishment of disturbance constraints. The disturbance constraint conditions include explicit and implicit disturbances. The explicit disturbance factors include product line changes, equipment failures, material shortages, and product quality non-conformities. The implicit disturbance factors include encompassing work-hour fluctuations and deviations from the workpiece completion time. Constraints caused by abnormal factors are established by varying the impact of multiple uncertain disturbances on the available production capacity time of the processes in the production system. Hence, this paper presents a mathematical modeling roadmap for workshop capacity driven by digital twins, which is shown in Figure 2. By clarifying the capacity reduction caused by disturbance factors, a mathematical model for discrete manufacturing workshop capacity is constructed, and prediction algorithms are employed to predict changes in workshop capacity.
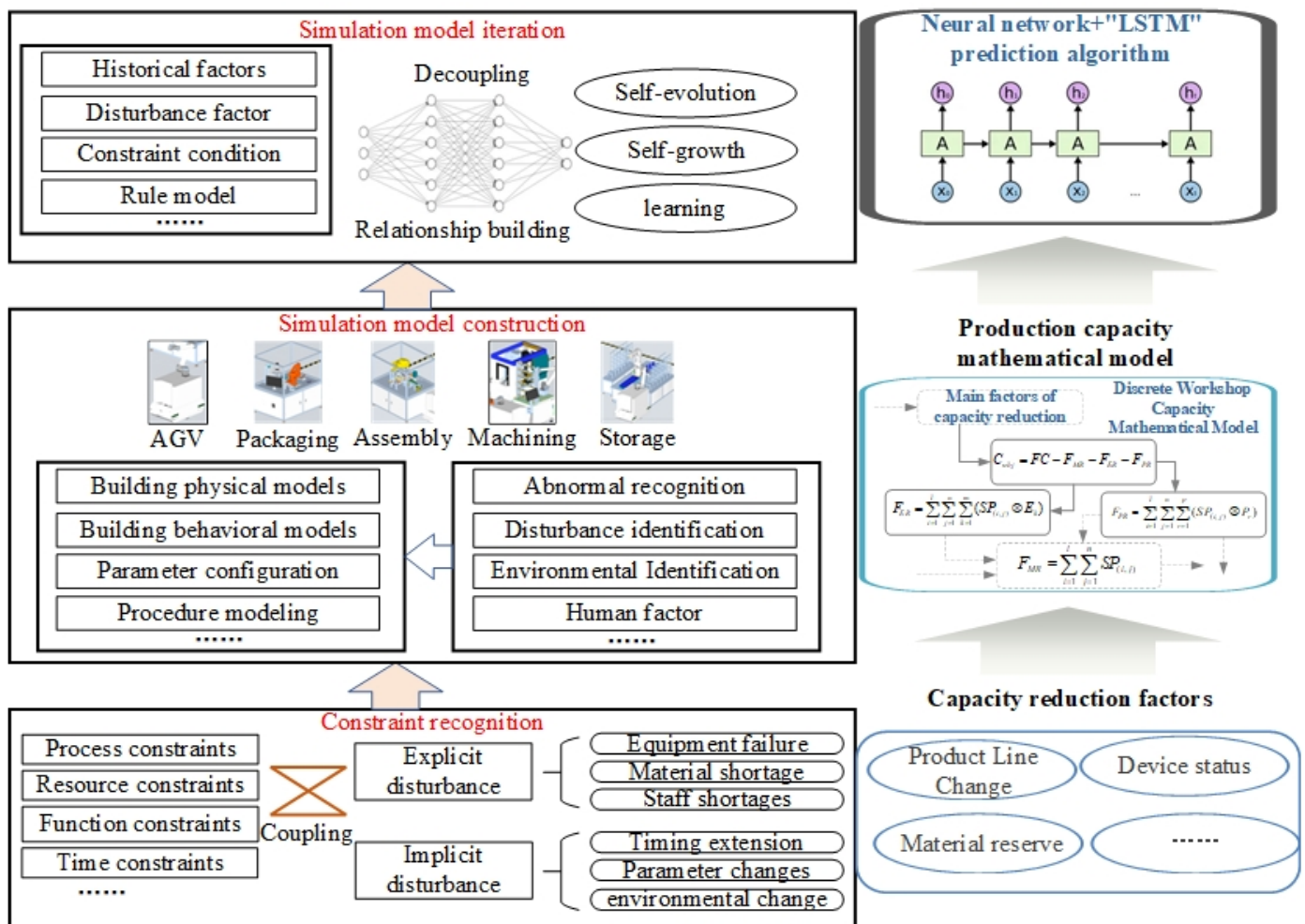

Figure 2. Mathematical modeling of workshop capacity driven by digital twins.

Capacity loss involves the identification of production bottlenecks, which can be attributed to various factors, such as delays in work-in-progress flow, process planning, process delays, material storage conditions, and equipment utilization. In a customized environment, the workshop requires considering the priority of orders and the urgency of

delivery times. The workshop digital twin system pre-simulates the disturbance factors that affect capacity loss through the previously collected workshop data and calculates the identification and transfer of capacity loss or production bottlenecks caused by uncertain disturbance factors. The production capacity model of a discrete workshop is a typical multi-constraint function constrained by production bottlenecks, as shown in Equation (1).

$$C_{obj} = FC - F_{MR} - F_{ER} - F_{PR} \tag{1}$$

where F_{ER} represents the production capacity reduced by insufficient equipment utilization, F_{PR} indicates the production capacity reduced by process fluctuations, and C_{obj} denotes the actual production capacity of the workshop.

Material shortages are an important factor in reducing workshop production capacity. For each product, the material requirements of the product, F_{MR}, are determined by the product type and product process, as shown in Equation (2):

$$F_{MR} = \sum_{i=1}^{l} \sum_{j=1}^{n} SP_{(i,j)} \tag{2}$$

where $SP_{(i,j)}$ represents the finished product caused by missing materials, l is the type of product in the workshop, and n symbolizes the number of processes in the finished product.

Equipment utilization rate, which refers to the ratio of the number of pieces of equipment involved in actual workshop production to the total number of pieces of equipment, is mainly influenced by factors such as the operational status of the equipment, the product types, and the process requirements and can be determined by Equation (3), as follows:

$$F_{ER} = \sum_{i=1}^{l} \sum_{j=1}^{n} \sum_{k=1}^{m} \left(SP_{(i,j)} \otimes E_k \right) \tag{3}$$

where E_{ij} represents the set of equipment to be maintained in the j process of workpiece j_i.

Changes in orders mainly cause a reduction in production capacity caused by process fluctuations. The same category, different models, and different batches of products often exhibit differences, which can introduce fluctuations in the pace of processing changes in the previous process and the connection between the subsequent processes during process planning. These fluctuations, in turn, result in a reduction in workshop production capacity. P_r represents the interval between the pre- and post-process segments where there is no need to wait and production is blocked. In this part, $P_r \subset \{P_1, P_2, \ldots, P_R\}$. F_{PR} is shown in Equation (4), in which $\otimes$ represents the mapping between the finished product $SP_{(i,j)}$ and P_r.

$$F_{PR} = \sum_{i=1}^{l} \sum_{j=1}^{n} \sum_{r=1}^{p} \left(SP_{(i,j)} \otimes P_r \right) \tag{4}$$

The above equation provides the main three causes of capacity loss. In actual workshops, the occurrence of production capacity disturbance factors is random and not fixed. Thus, the production system may be affected by multiple uncertain disturbance factors together, impacting the time variation in available production capacity. Let ΔT_{i1}, ΔT_{i2}, ΔT_{i3} be the loss of available production capacity duration caused by the three production capacity disturbance factors on the i process, as shown in Equation (5).

$$\begin{cases} \Delta T_{i1} = L_i \times \left[\sum_{0 \le c \le n} T_{ia} + \sum_{0 \le d \le n} T_{ib} \right] + T_{ic} \times BL_i \\ \Delta T_{i2} = L_i \times T_{id} \times EF_i \\ \Delta T_{i3} = L_i \times \left(\sum_{0 \le e \le n} T_{ie} + \sum_{0 \le f \le n} T_{if} + \sum_{0 \le g \le n} T_{ig} \right) \end{cases} \tag{5}$$

where L_i is the i process of a certain L category product, T_{ia} represents the available production capacity time lost in the i process due to the delay in the flow of work in process in the previous process, T_{ib} indicates the available production capacity time for the i process due to production line blockage or work-in-process blockage, and T_{ic} denotes the rework time of the product at the i process. In addition, BL_i is the rework quantity of the product, T_{id} symbolizes the time from the occurrence of a fault in the i process to the repair of the fault, EF_i represents the number of failures, T_{ie} defines the loss of available production capacity time caused by insufficient material supply, T_{if} means the loss of available capacity time caused by material handling tool damage, and T_{ig} refers to the loss of available production capacity time caused by quality issues with the materials on the workstation and failure to replenish them in a timely manner.

Equation (6) defines coupling multiple disturbance factors with production capacity time loss as a link, showing the possible capacity reduction in C_{loss}^i, coupling function and capacity availability rate for the i process.

$$\begin{cases} C_{loss}^i = SUM(\Delta T_{i1}(F_{MR}), \Delta T_{i2}(F_{ER}), \Delta T_{i3}(F_{PR})) \\ \eta_i = (FC_i - C_{loss}^i)/FC_i \end{cases} \tag{6}$$

5. Algorithm Design for Capacity Forecasting Driven by Twin Data

As can be seen in the previous section, a digital twin system is established before constructing a new factory and processing new products based on the construction of high-fidelity simulation models and the integration of constraints such as production factors, production plans, and production process control. Through the digital twin system, various data can be collected to evaluate the production capacity of new factories and new product processing. This prediction method can effectively address problems such as inaccurate predictions and the inability to meet delivery dates due to insufficient historical data and personnel experience. Production capacity analysis based on discrete manufacturing workshops involves the prediction of future production capacity based on twin data operations. Since the digital twin data belong to time series data, after analysis, this paper uses the short-term memory network algorithm of recurrent neural networks (RNNs) to predict production capacity.

LSTM is a temporal recurrent neural network designed to solve the long-term dependency problem of general RNNs, which can address the problems of gradient vanishing and gradient explosion in RNNs. The unique design structure of LSTM makes it suitable for processing and predicting important events with very long intervals and delays in the time series. Recurrent neural networks have the form of repeated module chains of neural networks. In the LSTM chain structure, each module contains four network layers, as shown in Figure 3. Additionally, LSTM has three gates (information-forgetting gate, input gate, and output gate) to optimize neuronal information.

The output of the result can be expressed as Equation (7):

$$h_t = o_t \times tanh(C_t) \tag{7}$$

where o_t is used to determine the output gate information transmitted by the neurons. From the input h_{t-1} of the previous neuron and the raw input, x_t is calculated by the activation function and weight parameters. It can be expressed as shown in Equation (8):

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{8}$$

where W_o and b_o are weight parameters and paranoia parameters, respectively.

In Equation (7), $tan\,h$ is the activation function with an output range of $[-1, 1]$. Moreover, the value range of the derivative function is 0–1, which can effectively solve the

problem of gradient disappearance in the cyclic network. The formula can be expressed as shown in Equation (9):

$$\tan h(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{9}$$

In Equation (8), σ represents the activation function, which is a smooth step function and differentiable. It can convert any value between 0 and 1 for binary classification using Equation (10).

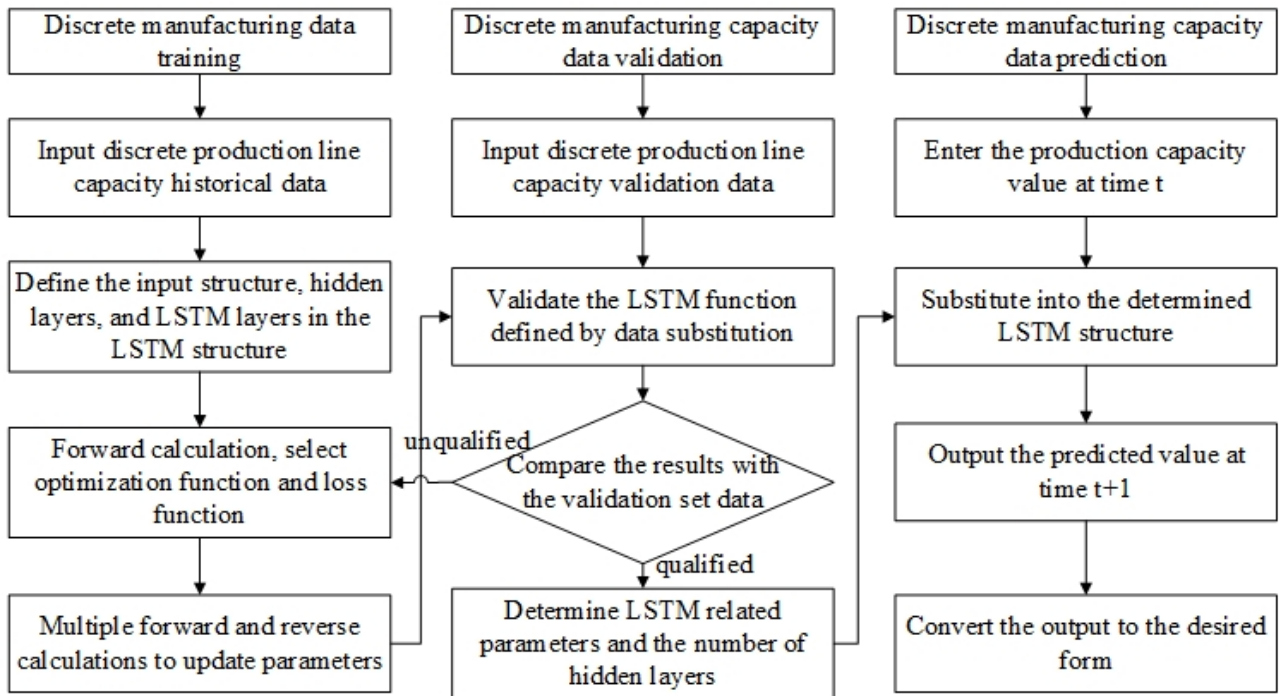$$\sigma = \frac{1}{1 + e^{-x}} \tag{10}$$

Discrete manufacturing data training	Discrete manufacturing capacity data validation	Discrete manufacturing capacity data prediction
Input discrete production line capacity historical data	Input discrete production line capacity validation data	Enter the production capacity value at time t
Define the input structure, hidden layers, and LSTM layers in the LSTM structure	Validate the LSTM function defined by data substitution	Substitute into the determined LSTM structure
Forward calculation, select optimization function and loss function	unqualified → Compare the results with the validation set data	Output the predicted value at time t+1
Multiple forward and reverse calculations to update parameters	qualified ↓ Determine LSTM related parameters and the number of hidden layers	Convert the output to the desired form

Figure 3. LSTM neural network processing data process.

C_t in Equation (7) represents the updated information of neurons, which effectively eliminates the long-term dependency problem in neural networks and establishes a more accurate connection between input and output information. The parameter C_t can be expressed as shown in Equation (11):

$$C_t = f_t \times C_{t-1} + i_t \times C_t' \tag{11}$$

where f_t indicates the forgetting gate information, which determines the information that must be discarded in the network state. In a sequence of data, different data may change in focus over time, forgetting unnecessary and unimportant information, which is also the reason for the emergence of forgetting gates. f_t can be expressed as shown in Equation (12):

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \tag{12}$$

where i_t and C_t' are input gate information. In the process of neuron update, the first step is to delete information, followed by the update of information. The network update information is obtained by transforming the previous neuron input h_{t-1} and system input x_t through the activation function of σ and $\tan h$. i_t can be expressed as shown in Equation (13):

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{13}$$

where C'_t can be expressed as shown in Equation (14):

$$C'_t = tanh\left(W_c\left[h_{(t-1)}, x_t\right] + b_c\right) \tag{14}$$

ΔT_{i1}, ΔT_{i2}, and ΔT_{i3} are taken as the input signals and propagate forward along the network. Under the action of each neuron weight, w_{ij}, and the activation function of $g(e)$, the output value of the network coupled with three disturbance factors is obtained through the output layer.

The input time series samples are considered as $X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ \dots & \dots & \dots \\ x_{t1} & x_{t2} & x_{t3} \end{bmatrix}$ in the input layer, in which X is the three disturbance factors that affect the workshop production capacity. The output value is subtracted from the actual value to obtain network error b. This error is transmitted backward through the neural network, and the connection weights between neurons in each layer are corrected layer by layer.

The hidden layers in the model are represented as shown in Equation (15):

$$h_{tj} = \sigma(\sum_{i=1}^{n} x_{ti}u_{ji} + h_{t-1,j}w_{jj}) \tag{15}$$

where x_{ti} is the value of the i node in the input layer at time t, and u_{ji} indicates the weight coefficient matrix between the j node in the hidden layer and the i node in the input layer. In addition, h_{tj} represents the output value of the j node of the hidden layer at time t, $h_{t-1,j}$ defines the output value of the j node of the hidden layer at time $t-1$, and w_{jj} symbolizes the weight coefficient matrix of the j node of the hidden layer during backpropagation.

The number of hidden layer nodes in the model is set to m, and its output value is $H = \begin{bmatrix} h_{11} & h_{12} & h_{1m} \\ \dots & \dots & \dots \\ h_{t1} & x_{t2} & x_{tm} \end{bmatrix}$. The model output layer is shown in Equation (16):

$$y_t = \sum_{j=1}^{m} \left(h_{tj}v_{kj}\right) + b \tag{16}$$

where y_t is the output value of the output layer at time t, V_{kj} indicates the weight coefficient matrix between the j node of the hidden layer and the k node of the output layer, and b denotes the bias term.

The output layer of the model is shown in Equation (17):

$$RMSE(y, y') = \sqrt{\sum_{i=1}^{n} \frac{(y_i - y'_i)^2}{n}} \tag{17}$$

where y represents the actual workshop production capacity, y' indicates the predicted value output by the model, and n denotes the number of samples. It mainly needs to determine the five hyperparameters of the model, namely the time steps of the input layer, the dimension of the input layer, the number of hidden layers, the dimension of each hidden layer, and the dimension of the output variable.

The LSTM neural network uses input information, x_t, and h_{t-1} to form the final output objective function h_t. By continuously iterating and updating the system, effective prediction results are formed. The iterative process of the LSTM neural network system is defined in Algorithm 1.

Algorithm 1: Algorithm for LSTM-based production capacity

    Input: Production capacity time series setting.
Output: Time series value of the production forecast.
    Begin
Initialize data by first-order difference
Data transformation into supervised learning
--------Offline model training process-------
Normalize the data set   //include ΔT_{i1}, ΔT_{i2}, ΔT_{i3}, and y
Training the LSTM_model
If $(RMSE(y, y') \leq \delta)$ **then**
    Export the LSTM_model
Else
    Adjusting parameters and re-fitting the LSTM_model
End if
-------Online model Test process-----------
For *i* **in range** (normalized test data)
    X= test_scaled[i, 0: -1]
    y= test_scaled[i, -1]
    Forecast_data= forecast_lstm(lstm_model, 1, X)
End for
Restore the first-order differential data
Store predicted values   //calculate the C_{loss}^{i} and η_i
Draw the comparison curve
End

---

6. Experimental Verification

The commemorative disk custom production line exemplifies a small, customized production line that belongs to a typical discrete manufacturing workshop. In the design process, facing the verification process of production line capacity is inevitable. As shown in Figure 4, the commemorative plate customization production line consists of seven modules, including lean material racks, grinding workstations, UV printing workstations, laser marking workstations, packaging workstations, storage workstations, and AGV workstations. The process is determined by the customer's order, where customers can choose a specific style of commemorative plate, text, and pattern according to their own needs. Using the RFID tag that facilitates the identification of binding information, the system determines the operational status of each workstation and verifies the completion of customized processing content. Once the processing is completed, all of these products enter the storage system and are subsequently transported to a fixed location by the AGV for delivery.

The twin system modeling approach for the commemorative disk custom production line based on twin data mainly includes three aspects. Among them, in terms of workshop element modeling, the three-dimensional geometric model based on CAD design, production process model, equipment motion model, and robot rule model is assembled and fused into workstation modules, such as storage workstations and assembly workstations. The workshop production line model is formed by assembling the workstation model. Among these models, the geometry, behavior, and rules of equipment and production lines are described. The production process modeling includes collected order data from the order system, actual production data from the simulation system, equipment processing efficiency data, and quality inspection data. According to the processing and fusion of various data, multi-dimensional analysis and optimization of production status can be realized. In terms of constructing the production system model, the complete production process of the commemorative disk custom production line is constructed based on the twin data and the virtual model, which provides effective support for subsequent production process data monitoring and visual analysis.
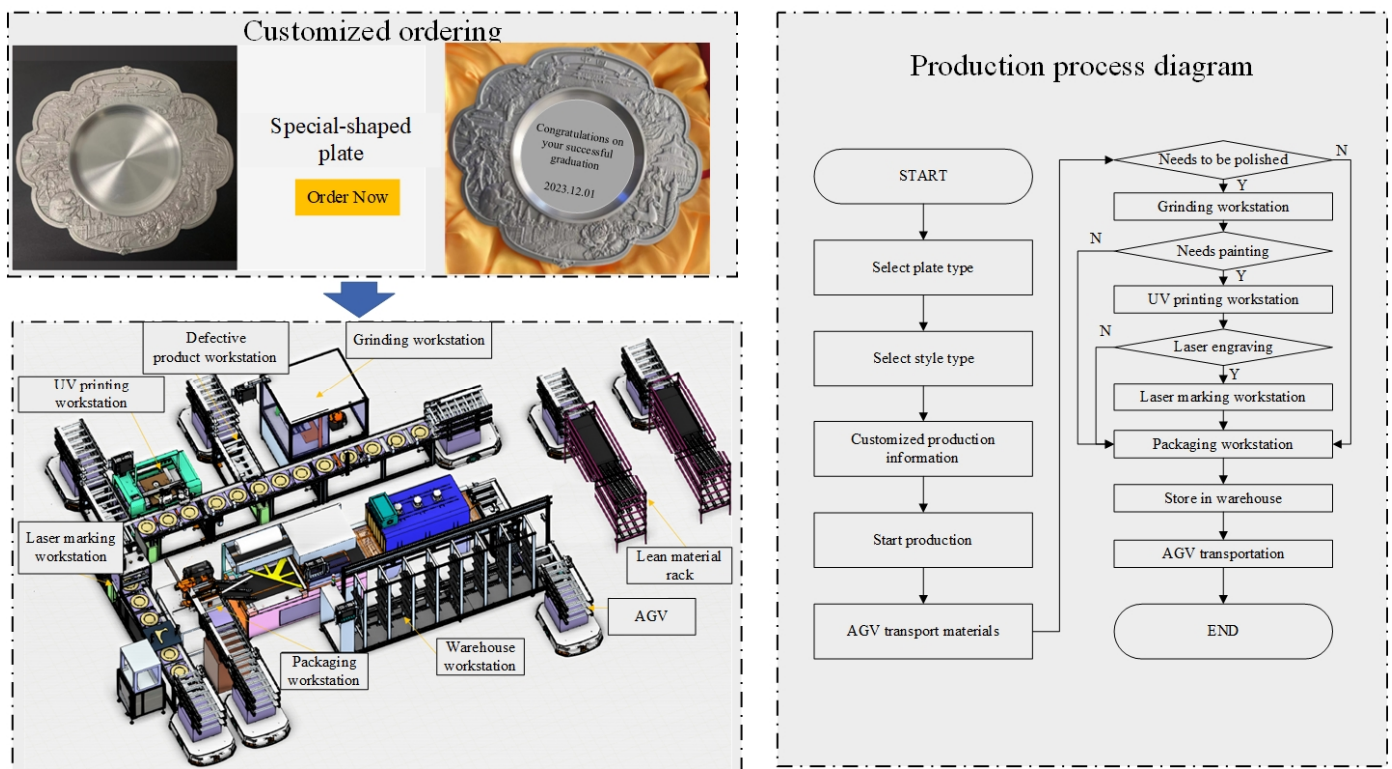
Figure 4. Customized production process of discrete workshop.

Since the production capacity is time series data, according to the above twin system model construction, the data are counted for one hour for production capacity statistics, as listed in Table 3.

Table 3. Production capacity statistics.

Serial Number	Order Processing Time (h)	Product Processing Quantity	Average Completion Time (s)
1	5	23	782.5
2	5	29	620.5
3	5	27	666.5
4	5	27	666.5
……	……	……	……
144	5	14	1285.5

This paper utilizes the twin data collected from the production process simulation to realize the entire workshop production capacity prediction process through the LSTM algorithm. And this article depicts the production status analysis process based on the combination of virtual and real models, which can be seen in Figure 5. The parameters used in the algorithm include a time window size (N) of 60, a step size of 1, 128 memory units, and three hidden layers. The prediction results are shown in Figures 6–8 for 1000, 2000, and 5000 iterations on the model, with the number of neurons set to four.
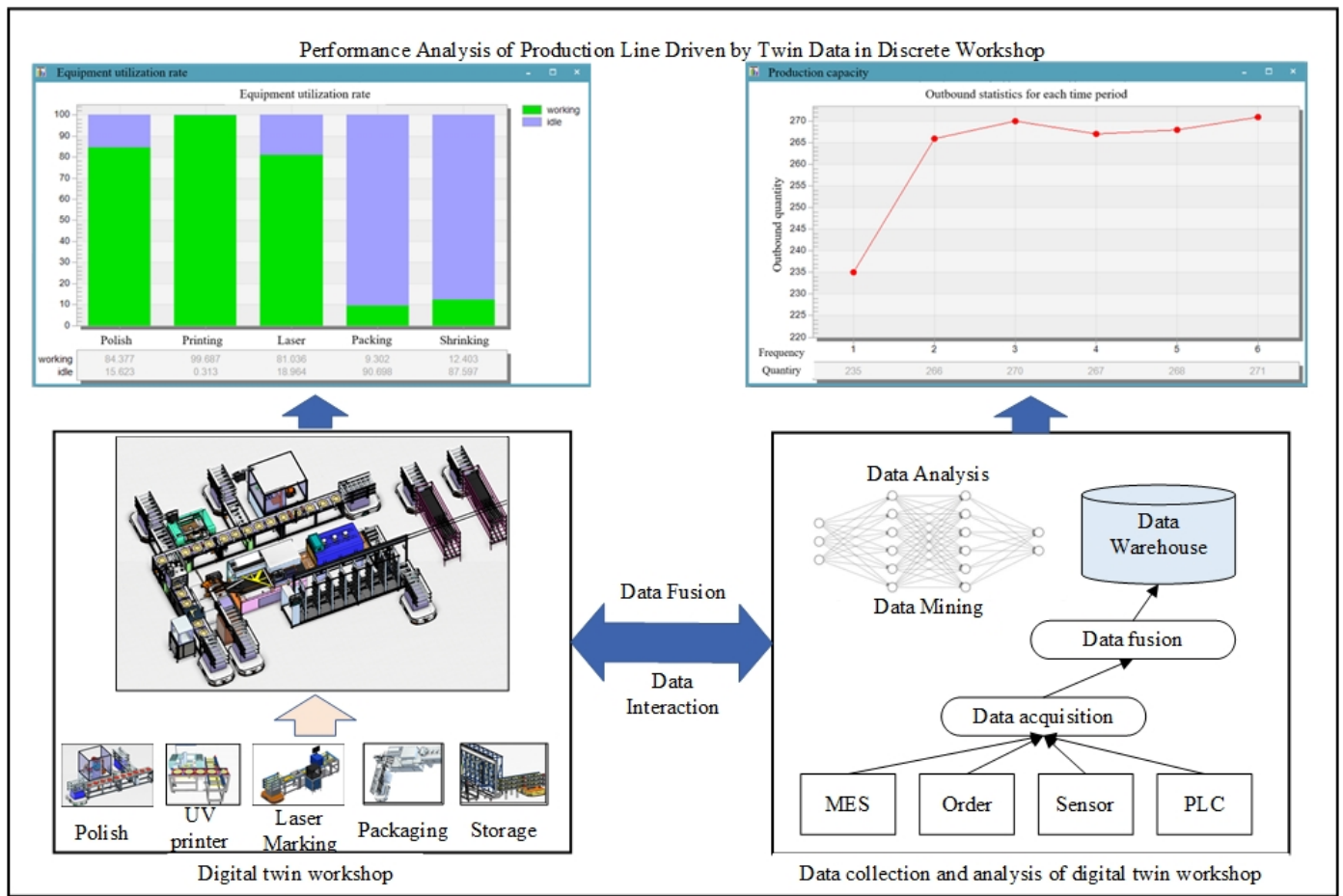
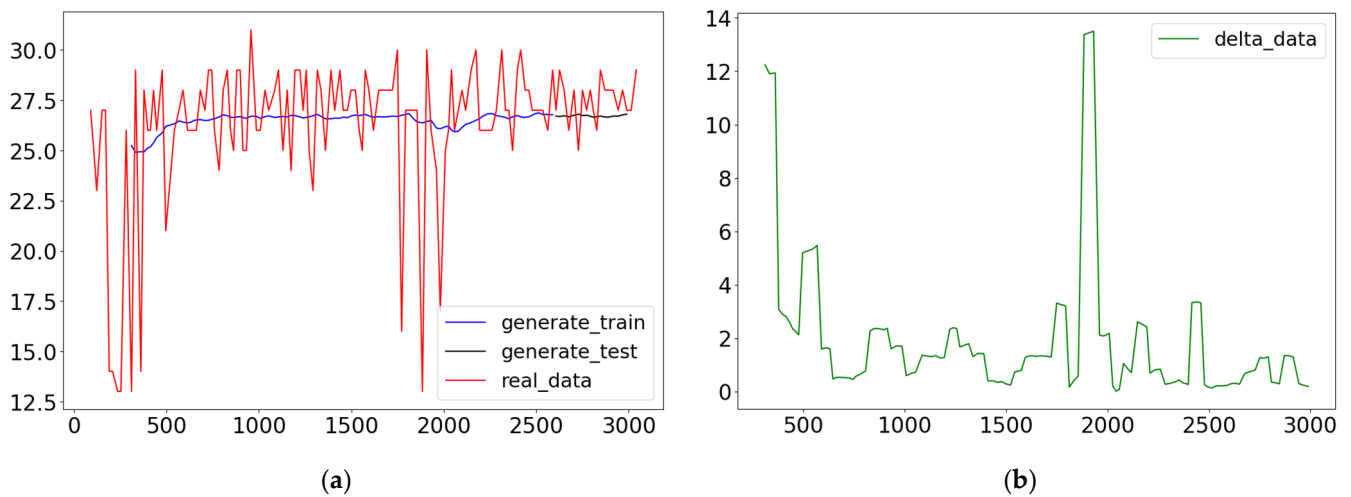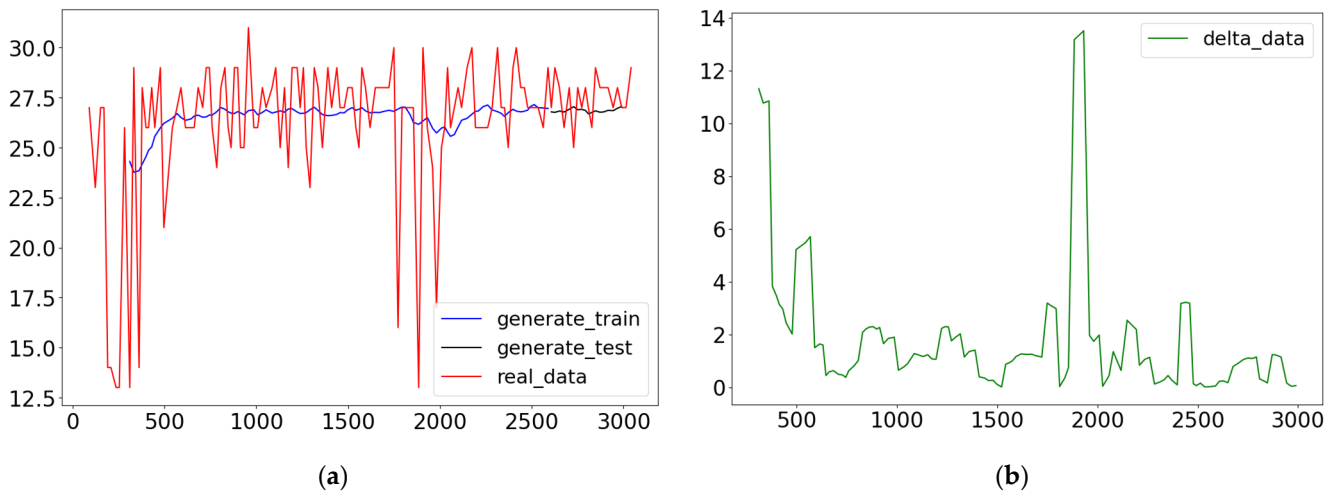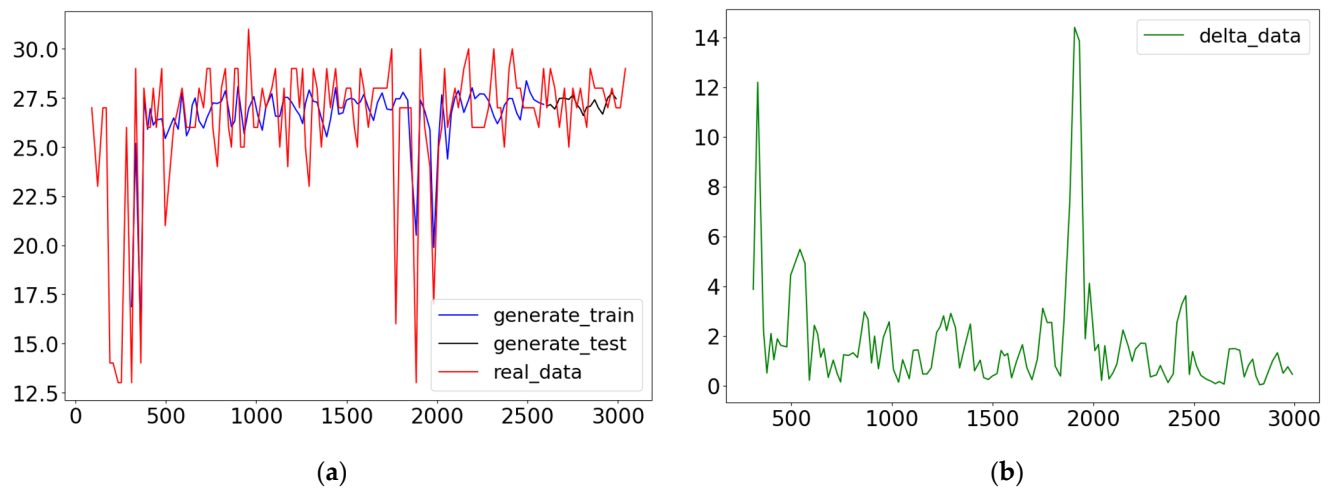**Figure 5.** Production status analysis driven by digital twin data.

(**a**)

(**b**)

Figure 6. LSTM production capacity prediction results (**a**) and variance (**b**) after 1000 iterations (1000/4).
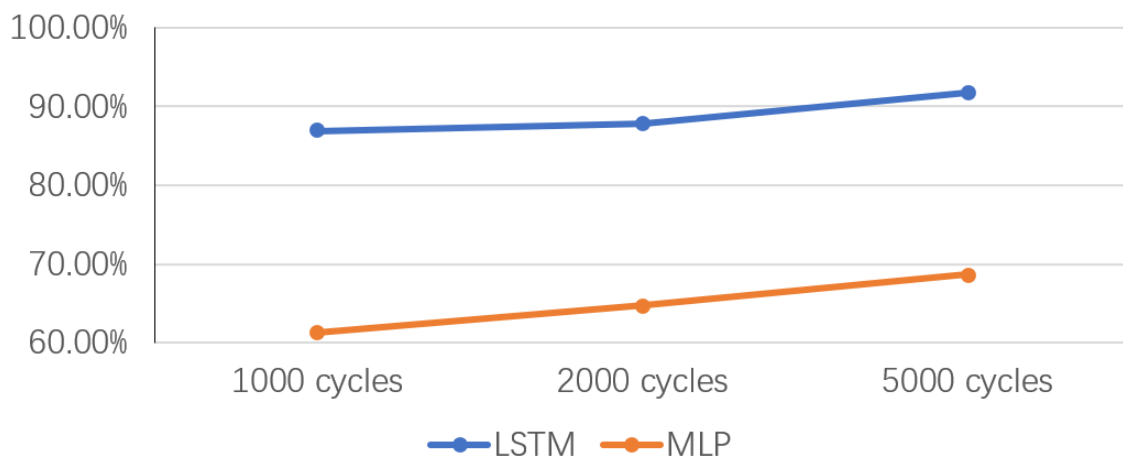
Figure 7. LSTM production capacity prediction results (**a**) and variance (**b**) after 2000 iterations (2000/4).

Figure 8. LSTM production capacity prediction results (**a**) and variance (**b**) after 5000 iterations (5000/4).

By calculating the algorithm results, the prediction accuracy of the algorithm is 0.869, 0.878, and 0.918, respectively. These results indicate that the system design conforms to the real performance and situation. Furthermore, as the number of iterations increases, the prediction accuracy gradually increases. At the same time, LSTM is a variant of traditional RNNs, which can effectively capture the semantic association between long sequences and alleviate the phenomenon of gradient disappearance or explosion compared with classic RNNs. In addition, the method is controlled by the gate structure, including the forget gate, input gate, cell state, and output gate, and its function is equivalent to adding a "processor" to judge whether the information is useful. It is also worth noting that the LSTM can better handle time series tasks and has better accuracy than a multi-layer perceptron (MLP: an artificial neural network that tends to structure). As shown in Figure 9, the LSTM has a better performance advantage than the MLP when the number of cycles is the same.

Figure 9. Comparison of LSTM and MLP prediction accuracy.

7. Conclusions

Before processing new products, the workshop's production capacity predictions are inaccurate, and the delivery date of new products often cannot be fulfilled due to the many types of products in discrete manufacturing workshops, small batches, and frequent production changes. To address the difficulty of predicting the production capacity of discrete manufacturing workshops, this paper analyzed the existing problems and related research in academia and proposed a digital twin-driven production capacity prediction framework for discrete manufacturing workshops. According to multiple disturbances, such as material shortages, equipment abnormalities, and process changes, a mathematical model for the production capacity reduction in discrete manufacturing workshops was given, and a production capacity prediction method of "twin data + LSTM algorithm" was innovatively proposed. Finally, the accuracy rate of production capacity prediction was verified using the commemorative disk custom production line as a prototype platform. The verification showed that the accuracy rate of the production line capacity prediction could reach 91.8%. The research in this paper was expected to predict the production capacity of new products from the current state and historical state of the workshop to establish a theoretical basis for evaluating product delivery. In addition, this paper only analyzed and discussed production capacity through the constructed digital twin model and did not explore deep-seated issues such as model evolution and iterative updates in the production process. Future research will investigate the evolution mechanism of the digital twin model for production processes, considering the iterative update of the twin data. Additionally, the impact of various elements of the production control process on workshop capacity will be further analyzed.

Author Contributions: Conceptualization, H.C. and J.W.; Methodology, H.C. and J.W.; Software, H.C. and B.C.; Validation, H.C. and B.C.; Resources, H.C.; Writing—original draft, H.C.; Writing—review & editing, J.W. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhou, J. Intelligent Manufacturing—Main Direction of "Made in China 2025". *China Mech. Eng.* **2015**, *26*, 2273.
2. Wang, S.; Wan, J.; Li, D.; Zhang, C. Implementing smart factory of industrie 4.0: An outlook. *Int. J. Distrib. Sens. Netw.* **2016**, *12*, 3159805. [CrossRef]
3. Tang, H.; Li, D.; Wan, J.; Imran, M.; Shoaib, M. A Reconfigurable Method for Intelligent Manufacturing Based on Industrial Cloud and Edge Intelligence. *IEEE Internet Things J.* **2021**, *7*, 4248–4259. [CrossRef]
4. Wang, X.; Hu, X.; Ren, Z.; Tian, T.; Wan, J. Knowledge-Graph Based Multi-Domain Model Integration Method for Digital-Twin Workshops. *Int. J. Adv. Manuf. Technol.* **2023**, *128*, 405–421. [CrossRef]
5. Leng, J.; Zhu, X.; Huang, Z.; Li, X.; Zheng, P.; Zhou, X.; Mourtzis, D.; Wang, B.; Qi, Q.; Shao, H.; et al. Unlocking the power of industrial artificial intelligence towards Industry 5.0: Insights, pathways, and challenges. *J. Manuf. Syst.* **2024**, *73*, 349–363. [CrossRef]
6. Wan, J.; Yin, B.; Li, D.; Celesti, A.; Tao, F.; Hua, Q. An Ontology-based Resource Reconfiguration Method for Manufacturing Cyber-Physical Systems. *IEEE ASME Trans. Mechatron.* **2018**, *23*, 2537–2546. [CrossRef]
7. Kong, T.; Hu, T.; Zhou, T.; Ye, Y. Data Construction Method for the Applications of Workshop Digital Twin System. *J. Manuf. Syst.* **2021**, *58*, 323–328. [CrossRef]
8. Zhang, Y.; Zhang, C.; Yan, J.; Yang, C.; Liu, Z. Rapid construction method of equipment model for discrete manufacturing digital twin workshop system. *Robot. Comput.-Integr. Manuf.* **2022**, *75*, 102309. [CrossRef]
9. Ren, Z.; Wan, J.; Deng, P. Machine-Learning-Driven Digital Twin for Lifecycle Management of Complex Equipment. *IEEE Trans. Emerg. Top. Comput.* **2022**, *10*, 9–22. [CrossRef]
10. Ozturk, O.; Begen, M.A.; Zaric, G.S. A branch and bound algorithm for scheduling unit size jobs on parallel batching machines to minimize makespan. *Int. J. Prod. Res.* **2017**, *55*, 1815–1831. [CrossRef]
11. Zhou, B.; Bao, J.; Li, J.; Lu, Y.; Liu, T.; Zhang, Q. A novel knowledge graph-based optimization approach for resource allocation in discrete manufacturing workshops. *Robot. Comput.-Integr. Manuf.* **2021**, *71*, 102160. [CrossRef]
12. Panetto, H.; Iung, B.; Ivanov, D.; Weichhart, G.; Wang, X. Challenges for the cyber-physical manufacturing enterprises of the future. *Annu. Rev. Control* **2019**, *47*, 200–213. [CrossRef]
13. Dengiz, B.; İç, Y.T.; Belgin, O. A meta-model based simulation optimization using hybrid simulation-analytical modeling to increase the productivity in automotive industry. *Math. Comput. Simul.* **2016**, *120*, 120–128. [CrossRef]
14. Benotsmane, R.; Kovács, G.; Dudás, L. Economic, social impacts and operation of smart factories in Industry 4.0 focusing on simulation and artificial intelligence of collaborating robots. *Soc. Sci.* **2019**, *8*, 143. [CrossRef]
15. Florescu, A.; Barabas, S.A. Modeling and simulation of a flexible manufacturing system—A basic component of industry 4.0. *Appl. Sci.* **2020**, *10*, 8300. [CrossRef]
16. Kim, S.; Ryu, K. Intelligent Process Quality Management for Supporting Collaboration of Mold Manufacturing SMEs. *Procedia Manuf.* **2020**, *51*, 381–387. [CrossRef]
17. Wang, W.; Zhang, Y.; Gu, J.; Wang, J. A proactive manufacturing resources assignment method based on production performance prediction for the smart factory. *IEEE Trans. Ind. Inform.* **2021**, *18*, 46–55. [CrossRef]
18. Zhang, B.; Li, X.; Wan, X. From digital twin to digital engineering modeling and simulation entering a new era. *J. Syst. Simul.* **2019**, *31*, 369–376.
19. Tao, F.; Qi, Q.; Wang, L.; Nee, A.Y.C. Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering* **2019**, *35*, 653–661. [CrossRef]
20. Lim, K.Y.H.; Zheng, P.; Chen, C.-H. A state-of-the-art survey of Digital Twin: Techniques, engineering product lifecycle management and business innovation perspectives. *J. Intell. Manuf.* **2020**, *31*, 1313–1337. [CrossRef]
21. Lu, Y.; Liu, C.; Kevin, I.; Wang, K.; Huang, H.; Xu, X. Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues. *Robot. Comput.-Integr. Manuf.* **2020**, *61*, 101837. [CrossRef]
22. Vachálek, J.; Bartalský, L.; Rovný, O.; Šišmišová, D.; Morháč, M.; Lokšík, M. The digital twin of an industrial production line within the industry 4.0 concept. In Proceedings of the 2017 21st International Conference on Process Control (PC), Strbske Pleso, Slovakia, 6–9 June 2017; pp. 258–262. [CrossRef]
23. Ma, J.; Chen, H.; Zhang, Y.; Guo, H.; Ren, Y.; Mo, R.; Liu, L. A digital twin-driven production management system for production workshop. *Int. J. Adv. Manuf. Technol.* **2020**, *110*, 1385–1397. [CrossRef]
24. Kunath, M.; Winkler, H. Integrating the Digital Twin of the manufacturing system into a decision support system for improving the order management process. *Procedia CIRP* **2018**, *72*, 225–231. [CrossRef]
25. Shumway, R.H.; Stoffer, D.S. ARIMA Models. *Springer Texts Stat.* **2011**, *90*, 3886–3896.
26. Liu, C.; Sun, B.; Zhang, C.; Li, F. A hybrid prediction model for residential electricity consumption using holt-winters and extreme learning machine. *Appl. Energy* **2020**, *275*, 115383. [CrossRef]
27. Bentéjac, C.; Csörgő, A.; Martínez-Muñoz, G. A comparative analysis of gradient boo Unlocking the Power of Industrial Artificial Intelligence towards Industry 5.0: Insights, Pathways, and Challengessting algorithms. *Artif. Intell. Rev.* **2021**, *54*, 1937–1967. [CrossRef]
28. Bullinaria, J.A. Recurrent neural networks. *Neural Comput. Lect.* **2013**, *12*, 1–20.

29.   Graves, A. Long short-term memory. *Neural Comput.* **2010**, *9*, 1735–1780.

30.   Jun, F.; Yang, L.; Zhang, R. Practical Privacy-preserving High-order Bi-Lanczos in Integrated Edge-Fog-Cloud Architecture for Cyber-Physical-Social Systems. *ACM Trans. Internet Technol.* **2019**, *19*, 1–18.