ISSN 2683-345X
journal homepage: http://ijiemjournal.uns.ac.rs/

International Journal of Industrial Engineering and Management

Volume 13 / No 2 / June 2022 / 119 - 134

Review article

Time series based forecasting methods in production systems: A systematic literature review

R. Hartner[a*] and V. Mezhuyev[a]

[a] *University of Applied Sciences FH JOANNEUM, Institute of Industrial Management, 8605 Kapfenberg, Austria*

ABSTRACT

Forecasting in production systems is used to anticipate their quality, efficiency, and yield. However, to the best of our knowledge, there exists no systematic review for industrial forecasting approaches. Thus, this work aimed to address this gap through a systematic literature review. The quantitative results revealed that industrial forecasting models are mainly applied in three economic sectors, with recurrent neural network models being the dominant approach. Moreover, this work proposes a classification of forecasting applications based on common characteristics found in reviewed sources. Several additional insights were produced, and future research directions were elaborated. Hence, this systematic review fosters an understanding of the current state-of-the-art of industrial forecasting approaches and facilitates future research initiatives.

ARTICLE INFO

Article history:

Received October 25, 2021
Revised April 4, 2022
Accepted May 5, 2022
Published online May 12, 2022

Keywords:
Industrial forecasting;
Machine learning;
Neural network;
Production system;
Systematic literature review

*Corresponding author:
Raphael Hartner
raphael.hartner2@fh-joanneum.at

1. Introduction

Digitalization and its use cases transform the industrial and production-oriented sectors worldwide [1-3]. In this regard, large-scale initiatives, such as Industry 4.0 and Made in China 2025, are concerned with the widespread adoption of advanced technologies and concepts, whereas several key research topics are commonly found throughout the literature [3, 4]. These topics are manifold and consist of the internet of things, cyber-physical systems, and big data applications, among others [5]. As a result of technological advancements in the context of digitalization, data are generated and available in an unprecedented magnitude and can be used to increase the efficiency of organizations [6].

In this regard, production-oriented companies employ data-driven approaches to a varying degree and in different contexts. Consequently, several (systematic) literature reviews were conducted to determine the state-of-the-art and future study directions to support researchers and practitioners in the production area, whereas two main objectives can be derived. On the one hand, existing systematic literature

Published by the University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia.

http://doi.org/10.24867/IJIEM-2022-2-306

reviews (SLRs) are focused on the applicability of general concepts, such as the usage of artificial intelligence (AI) and machine learning (ML) for digital twins in a wide range of industries [7] and potential applications for deep reinforcement learning [8] and machine learning [9] in manufacturing systems. The latter found that ML can be employed to increase efficiency in a wide area of applications, ranging from quality optimizations to fault diagnosis.

On the other hand, available literature reviews are conducted to identify state-of-the-art methods for specific applications or problems, such as predictive maintenance for industrial assets [10, 11], prediction and optimization of the drilling rate in oil/gas drillings [12], optimization of production processes [13] and mechanical fault diagnosis and prognosis in industrial manufacturing [14].

Even though manufacturing systems are usually characterized by non-linear and time-varying dependencies [14] and specific data-driven applications require not only a prediction of the current state, but also of future behavior [10, 11, 13], no dedicated review of state-of-the-art methods for forecasting in dynamical production systems exists to the best of our knowledge. Therefore, this work attempts to close this gap via a SLR focused on data-driven approaches for time series data to anticipate the performance of production systems, regarding quality, efficiency, and yield. Consequently, this paper is situated between SLRs focusing on the applicability of general concepts [7-9] and SLRs identifying state-of-the-art methods for specific applications [10-14], thus, complementing related work and contributing to the scientific discourse. Moreover, this SLR is not limited to production lines in the manufacturing domain but is concerned with production systems in general. Hence, other domains, for instance, oil or energy production are equally relevant for this SLR. This is due to the fact, that these systems are based on mechanical components and therefore have similar characteristics, which can be used to derive new information in a broader context. Furthermore, this paper is focused on ML models as well as statistical models.

The remaining part of this paper is structured as follows. In Section 2 the background of data-driven methods with different learning types and tasks are outlined. Section 3 describes the methodology of the SLR and its review protocol, whereas Section 4 evaluates literature sources and presents the results. Section 5 discusses the implications of the findings and Section 6 concludes this paper with future research directions.

2. Background of data-driven models

Before the SLR can be conducted, it is necessary to establish a common understanding of the research context. Different technologies and concepts, such as the internet of things and cyber-physical systems, facilitate the implementation of data-oriented applications and lead to significant opportunities for the industrial sector through data-driven techniques [6].

Generally, these data-driven techniques can be divided into *statistical* and *machine learning* models [12,15-18]. Statistical approaches are characterized by pre-selecting a model architecture for an investigated system, such as a linear regression with coefficients for each input feature [12, 15]. The data is used to estimate the parameters of the model to infer the relationships within the system [18]. Examples of statistical methods are linear regression, logistic regression, and principal component analysis [12, 16, 17].

In contrast to that, methods in the field of machine learning generally do not require a pre-selected model which determines the structure of the relationships but identify these aspects in an iterative learning procedure [18]. In other words, these approaches do not assume a priori specific mechanisms and structures within the investigated system. In that sense, machine learning is considered to be an algorithmic modeling approach [12, 15]. Examples of machine learning methods are decision tree, random forest, and support vector machine [16, 17]. Furthermore, even though there is no commonly used categorization of Bayesian techniques (statistical models [16], ML [17], separate Bayesian category [18]), the present SLR considers Bayesian techniques as ML methods, due to their iterative nature. However, although *neural networks* and deep learning represent a subfield for ML, its increasing popularity and success lead to a separate consideration to determine the current state-of-the-art of these specific approaches [11, 19].

2.1 Learning Types

Since the objectives of data-driven applications vary notably among different use cases, available algorithms focus on a diverse set of goals. Consequently, data-driven algorithms can be divided into four main learning types, as represented in Figure 1 [9, 14, 20, 21]. In this regard, *supervised learning* requires labeled data sets as input, thus, the target variable must be available, whereas the relationship between one or more features (independent variables) and the target variable (continuous or discrete) is learned [9, 16, 20].

In contrast to supervised learning approaches, *unsupervised learning* does not require labeled data sets. Thus, these types of algorithms cannot predict a continuous or a discrete value. However, unsupervised methods are used to identify patterns or extract features within a given data set without any prior knowledge of target values or dependencies among different attributes. In this regard, unsupervised learning algorithms are often deployed to determine similar data points (clustering), whereas the main difference in classification techniques is the absence of a target variable [20-22].

Even though supervised learning algorithms often lead to more accurate results in comparison to unsupervised learning approaches, the cost of labeling these data samples might outweigh the benefits, resulting in (partially) unlabeled data sets. To improve the performance of data-driven solutions, *semi-supervised approaches* are used to combine supervised and unsupervised methods [20, 21]. For instance, the concept of generative adversarial networks can be adapted to employ semi-supervised learning techniques, thus utilizing partially labeled data [23].

In addition to that, *reinforcement learning* represents another branch of data-driven learning. This technique differs significantly from previously discussed approaches since it is designed to interact with the environment and to consider its feedback [9, 21]. To achieve high-quality results, reinforcement learning takes real or simulated feedback into account and optimizes its predefined objective [24]. To find the global optimum, reinforcement methods need to explore different options, which can lead to a temporary decline in accuracy. In this regard, three main sub-categories of reinforcement learning exist [25]. These include dynamic programming, Monte Carlo methods and temporal difference. Importantly, the number of objectives in reinforcement learning systems is not limited to one. Thus, this learning type needs to find a compromise if multiple contradicting objectives are formulated. For this purpose, algorithms were introduced, which take not only immediate results, but also long-term consequences into account [26].

2.2 Tasks

With these learning types, several different tasks can be accomplished, whereas commonly found tasks are discussed in this section. First, the generic task of predicting a target (dependent) variable can be divided into *regression* (REG) and *classification* (CLASS) based on the target type [9, 13]. In the case of the former, the data-driven model predicts a continuous target variable, for instance, the length and width of the final product. On the other hand, classification is concerned with classifying data samples based on a discrete set of target classes – e.g., fault detection or image classification. Supervised learning is mainly used for both prediction variants [27].

Second, to determine clusters (groups) of similar samples and representational data points, the task of clustering is used as an unsupervised learning technique [10]. In this regard, two clustering types can be distinguished. While algorithms focused on *hard clustering* (HC) are concerned with assigning one specific cluster to each data sample, *soft clustering* (SC) algorithms allow multiple cluster assignments for each sample [28]. Consequently, the latter approach accounts for uncertainties during the clus-

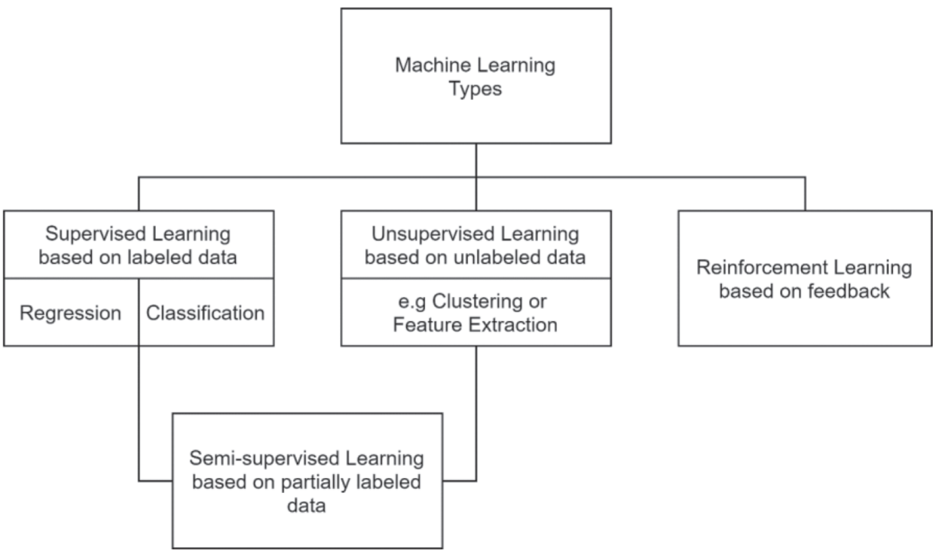

Figure 1. Main learning types for data-driven algorithms

tering process, which can be utilized in subsequent steps, for instance classifying new data points.

Third, to improve the quality of the data set and to reduce computational efforts, dimensionality reduction is often applied as a preprocessing task [9, 16, 29]. Thus, the main purpose is to extract (derive) new features or to select a subset of existing features (independent variables) from the data to optimize subsequent modeling steps. Consequently, the main difference between *feature extraction* (FE, also feature engineering) and *feature selection* (FS) is that the former calculates new features as a replacement for existing variables, whereas the latter is focused on selecting the optimal subset of the original feature set [29, 30]. To increase the robustness of FS, several selectors (homogeneous or heterogeneous) can be combined to form an ensemble selector [30]. Both tasks, FE and FS, can be accomplished by supervised as well as unsupervised learning techniques. The next section elaborates on the methodology and the review protocol of the SLR in this paper.

3. Methodology

To determine the state-of-the-art of industrial forecasting models, a SLR was conducted. A SLR is a specific type of literature review to answer predefined research questions in a systematic manner [31]. Consequently, the SLR results in reproducible findings which are based on a comprehensive set of available literature. For this purpose, research questions were formulated and the search process, as well as the search terms, were outlined in detail. Afterwards, literature sources were retrieved from online catalogs. The next steps were concerned with assessing the relevance based on the screening of the titles and abstracts and with removing duplicated entries. After that, a full-text assessment was performed with a set of specific exclusion criteria. The last step of the SLR was focused on extracting relevant data fields to answer the research questions.

3.1 Literature review planning protocol

To assess the state-of-the-art of industrial forecasting methods, the following research questions were defined for the SLR:

- RQ1: Which economic sectors are employing industrial forecasting models?
- RQ2: Which applications are addressed with industrial forecasting models?
- RQ3: Which data-driven categories (statistical, traditional ML, NN) are employed for forecasting in the industrial context?
- RQ4: Which data-driven categories are employed for each identified industrial forecasting application?
- RQ5: Which learning types are utilized in forecasting models?
- RQ6: Which tasks are performed in industrial forecasting scenarios?
- RQ7: Which algorithms are utilized for forecasting in the industrial context?

The following inclusion criteria were used to refine the search process in online catalogs and were specified to yield publications with high actuality:

- I1: Publication year after 2017 (year > 2017).
- I2: Publication type is journal paper, conference proceeding or book.
- I3: Publication language is English.
- I4: Full text is accessible in considered online catalogs.

The exclusion criteria were subsequently incorporated to build the specific search query and to guide the initial and full-text assessment:

- E1: Not related to an industrial context.
- E2: Does not consider time series data.
- E3: Does not include forecasting models.

After a relevant subset of literature sources was identified through the search procedure, the following data fields were extracted from each entry:

- F1: Domain of application.
- F2: Industrial forecasting application.
- F3: Data-driven category.
  (statistical models, traditional ML models, neural networks models).
- F4: Types of applied algorithms.
- F5: Learning type of considered models. (unsupervised, supervised, semi-supervised, reinforcement).
- F6: Performed tasks in industrial forecasting. (e.g. Clustering or Feature Extraction).

3.2 Search process

The review planning protocol serves as a foundation for the search process and ensures reproducible findings. To find state-of-the-art literature sources, the search process focuses on the online catalogs of IEEE Xplore and ScienceDirect. Both databases were queried on May 22, 2020. Corresponding categories were selected in ScienceDirect (review articles, research articles, book chapters) and IEEE Xplore (books, conferences, journals) to account for the inclusion criteria I2. Furthermore, the database search was conducted in the fields of title, abstract and author(-specific) keywords to determine relevant entries. The specific search query for both online catalogs consisted of four sub-terms, which represent individual aspects of the research questions to increase the relevance:

Time series AND (manufacturing OR production)
AND
(quality OR performance OR yield OR efficiency)
AND (forecast OR prediction)

These sub-terms are combined with AND clauses, whereas one or more keywords are included in each sub-term, which are found to be interchangeably used in the literature. In this regard, the third term is particularly important, since all four keywords are used in an indistinguishable manner to describe the performance, quality, efficiency, and yield of production systems. Furthermore, although there is a distinction between forecast and prediction, both are found synonymously used in related publications.

A summarizing overview and the number of publications at each step of the search process are illustrated in Figure 2. A total of 183 relevant entries were found with the search query and application of the inclusion criteria. The actual relevance was initially assessed by screening the title and abstract. This reduced the number of considered sources by 145, whereas the remaining 38 were checked for redundancy. However, no duplicates were found, and all 38 sources were included in the full-text assessment to determine the final relevance in a second iteration based on the exclusion criteria. The full-text assessment resulted in nine excluded literature sources, due to a lack of relevance (see Appendix A for details). Consequently, *29 conference and journal papers* (one relevant book chapter was excluded during the full-text assessment) were considered as input for the subsequent review and used to determine state-of-the-art approaches for industrial forecasting models.

4. Results of the systematic literature review

The systematic search process resulted in 29 literature sources, which were reviewed, and relevant data fields were extracted. Based on these data, the following sections represent the results of the SLR and consequently the answers to the research questions in a compressed form. Additionally, extracted details for each article are provided in Table 1.

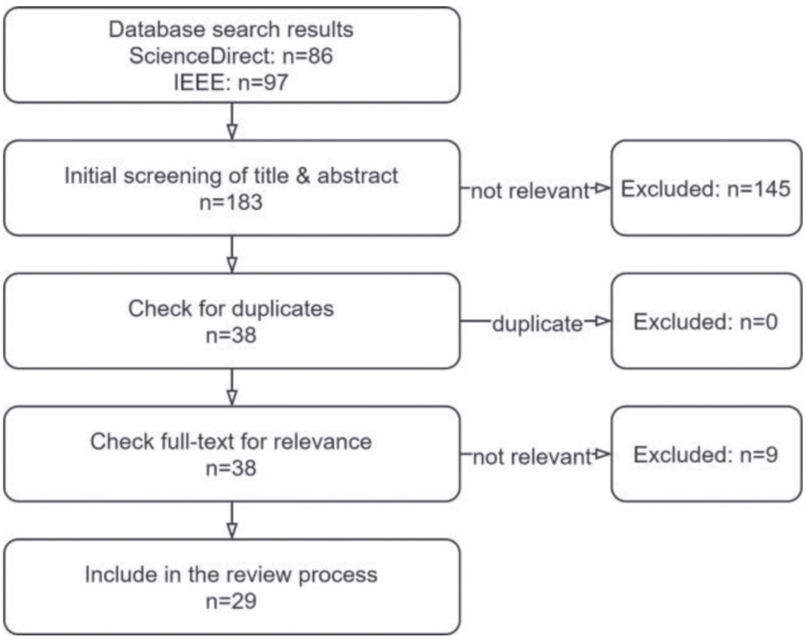

Figure 2. Process of the systematic literature review

Table 1. Extracted data fields of reviewed literature sources

Ref.	F1 Domain	F2 Application	F3 Category	F4 Algorithm	F5 Learning Type	F6 Task
[32]	D 35.11	Yield Prediction, Fault Prediction	NN Models	LSTM	Supervised	REG, CLASS
[33]	C 25.50	Quality Prediction	NN Models	MLP AE	Semi-Supervised	FE, REG, CLASS
[34]	C 23.51	Quality Prediction	Trad. ML Models	Fuzzy C-means, SVM	Semi-Supervised	SC, REG
[35]	C 29.10	Quality Prediction, Fault Prediction	NN Models	LSTM, B-LSTM	Supervised	REG, CLASS
[36]	D 35.11	Fault Prediction	Statistical Models	ARIMA	Supervised	REG
[37]	C 26.11	Quality Prediction, Process Optimization	Trad. ML Models	PLS, T-S fuzzy	Supervised	FE, REG
[38]	C 24.10	Quality Prediction	NN Models	LSTM	Supervised	REG
[39]	B 06.10	Process Behavior Prediction	Statistical Models	LR	Supervised	REG
[40]	C 26.11	Quality Prediction	Trad. ML Models	PLS, T-S fuzzy	Supervised	FE, REG
[41]	C 25.50	Fault Prediction, Quality Prediction	Statistical Models	PCA, LR	Semi-Supervised	FE, REG
[42]	C 24.42	Process Behavior Prediction	Trad. ML Models, NN Models	MLP, RF, RNN, LSTM, B-RNN, B-LSTM, CNN+MLP	Supervised	REG
[43]	B 06.10	Yield Prediction	NN Models	LSTM	Supervised	REG
[44]	C 26.51	Quality Prediction	NN Models	MLP AE	Semi-Supervised	FE, REG
[45]	C 24.10	Quality Prediction	Trad. ML Models, NN Models	RF-RFE, LSTM	Supervised	FS, REG
[46]	H 51.10, H 52.21	Fault Prediction	NN Models	SLP, B-LSTM	Supervised	REG
[47]	C 24.10	Quality Prediction	Statistical Models, Trad. ML Models	RR, RF, GBT	Supervised	FS, REG
[48]	C 24.10	Process Behavior Prediction	Trad. ML Models	T-S fuzzy	Supervised	REG
[49]	C 19.20, C 21.10	Quality Prediction	NN Models	LSTM	Supervised	REG
[50]	C 24.10	Process Behavior Prediction, Process Optimization	Statistical Models	Holt-Winters, ARIMA, Heuristic, MA	Supervised	REG
[51]	B 05.10	Process Behavior Prediction	NN Models	GRU	Supervised	REG
[52]	C 26.11	Quality Prediction	Trad. ML Models	PLS, T-S fuzzy	Supervised	FE, REG
[53]	C 26.11	Process Behavior Prediction	NN Models	B-GRU AE	Semi-Supervised	FE, REG
[19]	D 35.11	Fault Prediction	NN Models	CNN	Supervised	CLASS
[54]	C 26.11	Process Behavior Prediction	Trad. ML Models, NN Models	MARS, RF, GBT, NB, K-NN, SVM, MLP, LSTM	Supervised	FS, CLASS
[55]	B 06.10	Yield Prediction	NN Models	LSTM	Supervised	REG
[56]	C 22.22, C 27.90	Fault Prediction	Statistical Models	ARIMA	Supervised	REG
[57]	B 06.10	Yield Prediction	NN Models	LSTM	Supervised	REG
[58]	C 25.62	Fault Prediction	NN Models	CNN	Supervised	REG
[59]	C 13.10	Process Behavior Prediction	NN Models	GRU	Supervised	REG

RQ1: Which economic sectors are employing industrial forecasting models?

To classify the economic sector in which industrial forecasting models are applied, the *Statistical Classification of Economic Activities in the European Community* is used [60]. This classification is hierarchically structured and uses four levels. The first level refers to the general domain (e.g. "manufacturing" or "mining and quarrying"), whereas the fourth level of classification specifically describes the economic activity, such as "production of electricity" or "manufacture of electronic components". To associate reviewed literature sources with one or more economic activities, the context of the described applications was used. As a result, four domains of economic activities applied industrial forecasting models, as summarized in Table 2. A total of five literature sources validated their models in the "mining and quarrying" (B) sector, 22 sources in "manufacturing" (C), three sources in electricity, gas, steam and air conditioning supply" (D) and two sources in the "transportation and storage" (H) sector. Since three literature sources considered two sectors, the total number amounts to more than 100 % in comparison to the number of reviewed sources. In one case, the validation data set was concerned with an aircraft engine, which is applicable for both sectors, passenger, and freight air transportation (H 51.10 and H 52.21). As shown in Table 2, three sectors (B 06.10, C 24.10, C 26.11) are considered in 43.8 % of all 32 sector references.

RQ2: Which applications are addressed with industrial forecasting models?

Although industrial forecasting applications are individually designed for a specific purpose, common characteristics are found in reviewed literature sources. Consequently, five generic applications were identified in this review:

- *Yield prediction*: In complex and non-linear environments, such as energy and oil production, the future output (yield) of the process cannot be easily anticipated. Therefore, yield prediction based on time series data is applied to forecast the process' output.
- *Fault prediction*: The prevention of upcoming faults and defects in industrial machinery is essential for efficient processes. In this regard, the prediction of the remaining useful life of machines and tools as well as anomaly detection are both relevant subfields. Thus, fault prediction is used to forecast future issues and enable machine operators or control units to proactively prevent or mitigate faults and deviations.

Table 2. Number of references to each economic sector

Economic sector	References	
B 05.10 - Mining of hard coal	1	5
B 06.10 - Extraction of crude petroleum	4	
C 13.10 - Preparation and spinning of textile fibres	1	22
C 19.20 - Manufacture of refined petroleum products	1	
C 21.10 - Manufacture of basic pharmaceutical products	1	
C 22.22 - Manufacture of plastic packing goods	1	
C 23.51 - Manufacture of cement	1	
C 24.10 - Manufacture of basic iron and steel and of ferro-alloys	5	
C 24.42 - Aluminium production	1	
C 25.50 - Forging, pressing, stamping and roll-forming of metal; powder metallurgy	2	
C 25.62 - Machining	1	
C 26.11 - Manufacture of electronic components	5	
C 26.51 - Manufacture of instruments and appliances for measuring, testing and navigation	1	
C 27.90 - Manufacture of other electrical equipment	1	
C 29.10 - Manufacture of motor vehicles	1	
D 35.11 - Production of electricity	3	3
H 51.10 - Passenger air transport	1	2
H 52.21 - Freight air transport	1	

- *Quality prediction*: The product quality (e.g., chemical properties) in industrial processes is often influenced by many factors, which are not always controllable, for instance, environmental aspects (e.g., temperature, humidity). Correspondingly, quality predictions are used to prevent quality deviations in advance through controllable factors.
- *Process behavior prediction*: Controlling a production process is not only required for consistent product quality, but also for safety and efficiency reasons. Hence, predicting the future behavior of processes (e.g., pipeline pressure, motor torque) is the basis for efficient and effective process control to mitigate deviations and faults.
- *Process optimization*: Complex and interrelated processes are difficult to control and optimize, since many influential factors must be considered. Consequently, process optimization, for instance, to improve the quality, based on industrial forecasting models are utilized to make use of untapped potentials within production processes.

As depicted in Figure 3, out of 29 reviewed literature sources 12 considered quality prediction as the application for their industrial forecasting models. Additionally, eight sources focused on fault prediction as well as on process behavior prediction. These three applications represent the majority of reviewed literature sources. On the other hand, yield prediction was considered by four literature sources and was exclusively applied to energy and oil production. Furthermore, only two sources applied forecasting models in the context of process optimization. Since five reviewed literature sources considered two applications, the sum of the paper counts results in more than 100 %.

RQ3: Which data-driven categories are employed for forecasting in the industrial context?

To implement a data-driven application, different modeling categories can be employed, whereas Figure 4 contains the distribution of data-driven categories in reviewed sources. As illustrated, NN models represent the majority of data-driven approaches with 62 % and traditional ML techniques rank second (31 %). Statistical models are employed by six sources to implement an industrial forecasting application (21 %). Since four considered literature sources applied two categories of data-driven approaches, the sum of the paper counts results in more than 100 %.

RQ4: Which data-driven categories are employed for each identified industrial forecasting application?

To derive new insights concerning applicable models, an analysis of data-driven categories and identified industrial forecasting applications was conducted. As summarized in Table 3, apart from process optimization, NN models are similarly distributed among the forecasting applications. However, process optimization was only considered by two reviewed literature sources in total, thus representing the minority of application types. On the other hand, traditional ML models were focused on quality prediction and process behavior prediction, whereas the remaining applications were not or only once considered. Additionally, statistical models were applied for fault prediction, process behavior prediction and process optimization. Since four reviewed sources applied two data-driven categories and five other sources considered two separate applications, a total of 38 pairs of application data-driven categories are included in 29 reviewed literature sources.

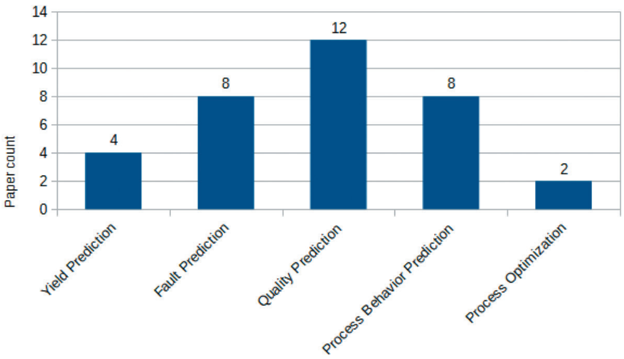

Figure 3. Paper count for each forecasting application

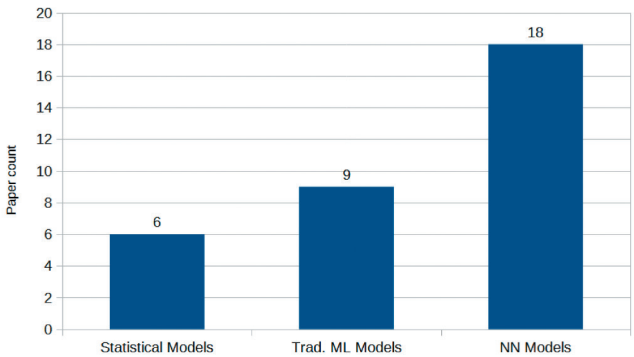

Figure 4. Paper count for each considered data-driven category

RQ5: Which learning types are utilized in forecasting models?

As shown in Figure 5, two learning types are applied for industrial forecasting models. Supervised learning is utilized by 83 % of reviewed literature sources. Even though five sources (17 %) implemented a semi-supervised learning approach with both unsupervised and supervised characteristics, no source was found with an exclusively unsupervised learning model. Moreover, reinforcement learning was not applied by any of the reviewed literature sources.

RQ6: Which tasks are performed in industrial forecasting scenarios?

Figure 6 depicts the distribution of data-driven tasks in industrial forecasting applications. As illustrated, most literature sources (93 %) developed models for *regression* tasks. On the other hand, *classification* is employed by five out of 29 literature sources (17 %), whereas three applied both, regression, and classification models. Other tasks, such as feature extraction (FE), feature selection (FS) and soft clustering (SC), are consistently performed in combination with either regression or classification. Depending on the specific type of algorithm, these combinations result in semi-supervised learning approaches (see above). Furthermore, none of the reviewed literature sources applied hard clustering (HC).

RQ7: Which algorithms are utilized for forecasting in the industrial context?

The analysis of applied algorithms in the context of industrial forecasting is divided into statistical, traditional ML and NN models (see above for the definition of data-driven categories). Since 12 out of 29 literature sources utilized two or more algorithms, the sum of the paper counts results in more than 100 %. First, as shown in Figure 7, seven statistical methods are employed in the considered literature. Three sources used autoregressive integrated moving average (ARIMA) models to forecast process behavior and faults. Additionally, Holt-Winters, moving average (MA) and a custom heuristic are applied in combination with ARIMA in one case. Linear regression (LR) is used in two references, whereas the related ridge regression (RR) is applied once. In this regard, RR utilizes L2 regularization to improve the robustness in comparison to conventional LR models [61]. Moreover, principal component analysis (PCA) is used by one source to extract linearly independent features for subsequent modeling processes.

Second, many different algorithms are used for traditional ML models, as depicted in Figure 8. In this regard, Takagi-Sugeno (T-S) fuzzy systems rank first (four sources) and three sources use partial least square (PLS) FE techniques before applying T-S fuzzy algorithms. Recursive feature elimination (RFE)

Table 3. Number of forecasting models for each data-driven category and application

Application	Statistical models	Traditional ML models	NN models
Yield Prediction	0	0	4
Fault Prediction	3	0	5
Quality Prediction	2	6	6
Process Behavior Prediction	2	3	5
Process Optimization	1	1	0

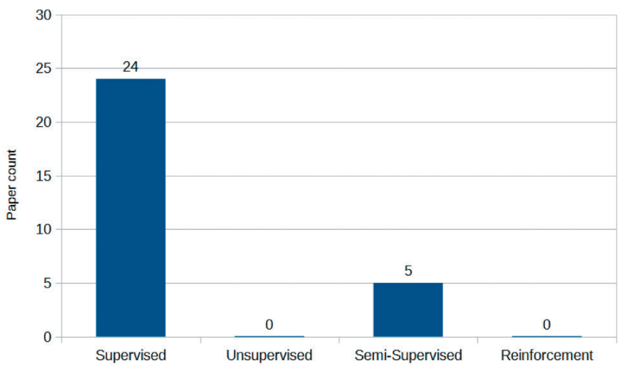

Figure 5. Paper count for each considered learning type

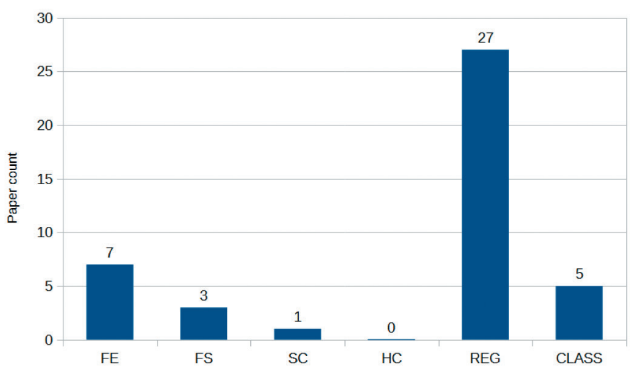

Figure 6. Paper count for each considered data-driven task

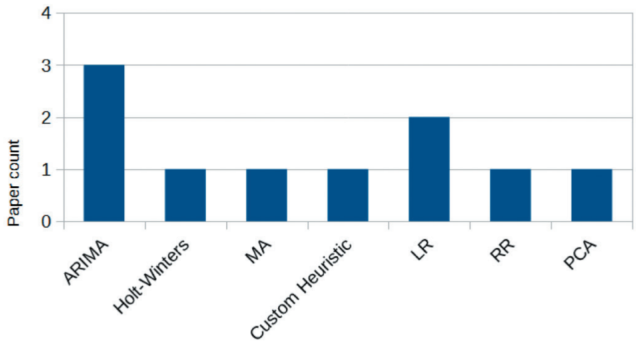

Figure 7. Paper count for each considered statistical algorithm

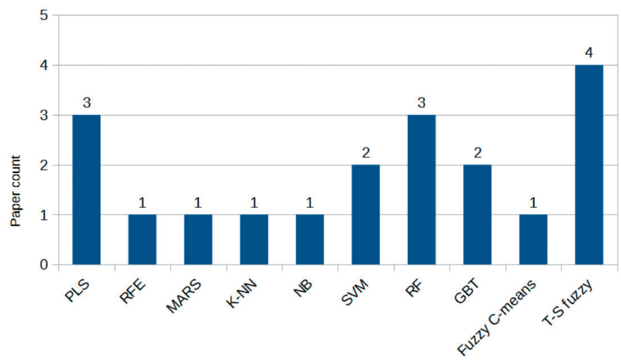

Figure 8. Paper count for each considered traditional ML algorithm

based on random forest and the multivariate adaptive regression splines (MARS) algorithm are used for FS by one literature source each. Support-vector machines (SVM) is mentioned in two references. Ensemble techniques, such as random forest (RF) and gradient boosting trees (GBT), are implemented by three and two sources respectively. Furthermore, one source proposed an ensemble technique based on support-vector machines, leading to 44 % of traditional ML sources using ensemble models. The algorithms k-nearest neighbor (K-NN) and naive Bayes (NB) are applied by one literature source each. Moreover, the fuzzy c-means algorithm represents the only clustering approach in considered sources.

Third, similar to traditional ML models, a large number of different NN algorithms are applied by corresponding literature sources, as illustrated in Figure 9. However, long short-term memory (LSTM) algorithms are utilized in ten cases (56 % of NN sources). Basic recurrent neural networks (RNN) and gated recurrent units (GRU) are employed by one and three sources respectively. Together with their bidirectional derivatives (B-LSTM, B-RNN, B-GRU-AE), recurrent approaches are used by the majority of reviewed NN sources for industrial forecasting scenarios. Nevertheless, other algorithms, such as convolutional neural networks (CNN) or feed-forward single-/multi-layer perceptrons (SLP/MLP), are also considered in the literature. As shown in 9, three sources implemented an auto-encoder (AE) in combination with MLP (MLP-AE) and B-GRU (B-GRU-AE), to extract compressed features, before conducting the forecast.

5. Discussion

Based on the *results*, this section provides a discussion for each research question, whereas findings and limitations are outlined.

RQ1. Based on the *Statistical Classification of Economic Activities in the European Community* [60], 18 economic sectors were identified in reviewed literature sources. In this regard, three sectors contributed a combined 43.8 % to all 32 sector references:

- B 06.10 - Extraction of crude petroleum (12.5 %)
- C 24.10 - Manufacture of basic iron and steel and of ferro-alloys (15.6 %)
- C 26.11 - Manufacture of electronic components (15.6 %)

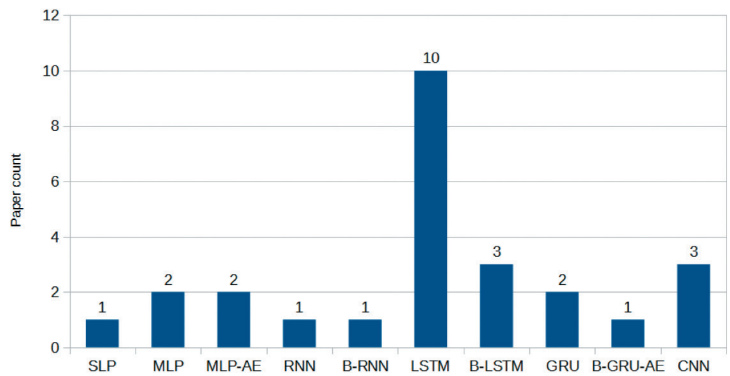

Figure 9. Paper count for each considered NN algorithm

As a result, it can be concluded that these three economic sectors are actively applying industrial forecasting models to optimize their operations. Consequently, future research can focus on transferring successful methods to other sectors, which do not yet rely on sophisticated industrial forecasting models.

RQ2. Even though the applications in reviewed literature sources are individually developed for specific requirements, common characteristics were identified and summarized in five generic applications. As revealed by the paper count, quality predictions are the main field of application for industrial forecasting models (35.3 %), whereas both applications, fault prediction and process behavior prediction, contribute 23.5 % to the total paper count. Consequently, the remaining applications in the context of yield prediction (11.8 %) and process optimization (5.9 %) can profit from methods applied in the three dominant applications. In particular, proactively optimizing a process in advance requires other forecasting models for relevant aspects, such as faults, quality and process behavior. As a result, technological and methodological progress in these applications benefits proactive process optimization. Regarding the generic production line problems, stated by [9], industrial forecasting applications based on time series data can contribute to these aspects, as shown in Table 4. However, forecasting applications are not only applied in the context of production lines but for production in general. Thus, other economic sectors, for instance, oil and energy production, benefit from these applications as well.

RQ3. In contrast to related SLR papers [9, 10], it was found, that NN algorithms are applied by the majority (62 %) of reviewed literature sources. Consequently, it can be concluded, that these algorithms are particularly well suited for forecasting models due to supporting non-linear relationships, which are often found in time series data. However, traditional ML algorithms, such as random forest or support vector machines, were employed by 9 out of 29 reviewed sources. Thus, representing a considerable alternative to NN approaches, especially when interpretable forecasting results are required due to the black box characteristics of NN models. Finally, only a minor portion of 21 % of reviewed literature sources applied statistical models. Hence, the current state-of-the-art of data-driven categories in industrial forecasting are NN approaches followed by traditional ML. In this regard, future research could focus on hybrid methods consisting of traditional ML and NN to incorporate the benefits of both data-driven categories.

RQ4. Yield and fault prediction are dominated by NN models, which indicates that the non-linear and complex modeling capabilities of neural networks are appropriate for these applications. However, depending on the characteristic of the data set and the individual goals, traditional ML models can also be considered, since these models were successfully employed for quality and process behavior prediction as well. In particular, if explainable results are required for the application, traditional ML models have an advantage in contrast to the black box characteristics of neural networks. Consequently, future research could focus on transferring effective methods from quality and process behavior prediction applications to yield and fault prediction. Furthermore, since statistical models are also applied in nearly all applications (apart from yield prediction), it can be concluded that the applicability mainly depends on the specific characteristic of the application. Particularly, if only linear dependencies are expected, statistical models are well suited for the task.

RQ5. In regard to the learning types, this SLR found that all reviewed literature sources contained supervised learning algorithms. However, five references additionally applied unsupervised learning, resulting in semi-supervised models to improve the results. Thus, incorporating both learning types, supervised and unsupervised, into industrial forecasting models and developing novel hybrid approaches can be considered to be subject to future work. More-

Table 4. Association of production line problems and industrial forecasting applications

Production line problems [9]	Industrial forecasting application
Quality optimization Product failure detection	Quality prediction
Fault diagnosis Preventive maintenance	Fault prediction Process behavior prediction
Scheduling optimization	Process behavior prediction Process optimization
Waste reduction	Process optimization
Yield improvement	Yield prediction

over, since none of the reviewed literature sources considered reinforcement learning, exploring this learning type might lead to new progress in the field of industrial forecasting. In particular, the application of process optimization (potentially with simulated assets) could benefit from reinforcement learning [62].

RQ6. The analysis of data-driven tasks revealed that all reviewed literature sources contained either regression or classification (or both) tasks. Even though most sources relied on regression-based forecasting models, it was shown that classification can also be utilized for forecasting. However, the target must be properly adapted to represent a classification problem. The remaining tasks (FE, FS, SC) are applied in conjunction with regression or classification to improve the results. Therefore, these tasks are not suited for stand-alone usage in forecasting models.

RQ7. Both, regression and classification tasks, are performed by NN models and the specific characteristics of time series data leads to a dominant role of recurrent NN models (RNN, LSTM, GRU) and their bidirectional derivatives. In particular, LSTM models are mainly applied in industrial forecasting applications (35 % of all reviewed sources). Auto-encoders based on NN are used to extract compressed features, which normally yield superior results. Nonetheless, these FE techniques are currently only applied by a minor subset of reviewed sources, hence, additional opportunities are expected in other applications. Although, CNN models are inherently well suited for image processing (e.g., object recognition), they can be employed for time series forecasts as well. For this purpose, the characteristic of CNN algorithms to detect patterns among neighboring data points can be exploited for time series data and enhanced in future research.

6. Conclusion

To sum up this work, data-driven models are employed across different industries and contribute to the *competitive* advantage of implementing organizations. These models are also utilized for industrial forecasting applications to anticipate the performance, quality, efficiency, and yield in production systems. However, due to time series consideration, forecasting conceptually differs from other data-driven models. Therefore, in contrast to existing literature, this work conducted a dedicated SLR to determine the state-of-the-art of industrial forecasting models.

Consequently, the systematic search process was based on specific research questions and search terms, resulting in an initial screening of 183 literature sources. During this first phase, 38 sources were found to be relevant. Nevertheless, after a full-text assessment was performed, 29 journal papers and conference proceedings were included in the review, where predefined data fields were extracted to answer the research questions.

As shown by the quantitative results, although several economic sectors are referenced in the literature, three were found to be particularly focused on industrial forecasting models. Additionally, this study identified five generic industrial forecasting applications based on common characteristics determined in the reviewed literature. In this regard, forecasting models are mainly applied for quality, fault, or process behavior prediction, as shown by the results. In contrast to related SLR studies, it was found, that NN models represent the majority of data-driven categories, whereas traditional ML, such as random forest, is also considered to be a relevant approach. Nevertheless, these two categories are not equally employed across different applications. As determined by this study, while NN models are similarly distributed among different applications, traditional ML models are mostly focused on quality and process behavior predictions. Additionally, the findings revealed that all forecasting models incorporate at least one supervised component. Importantly, only five sources considered a semi-supervised approach and no reference to reinforcement learning was found. In accordance with these findings, all reviewed sources focused on either regression or classification (or both) as supervised learning tasks. Other aspects, such as feature extraction, feature selection and soft clustering, are exclusively used in conjunction with these tasks. Furthermore, the industrial forecasting models are dominated by recurrent neural networks, in particular by the LSTM algorithm, and their bidirectional derivatives. Moreover, auto-encoders are suitable to extract compressed features, while CNN models are used to detect patterns among neighboring data points.

As stated in the introduction, with these findings, this study closes the research gap for forecasting applications in the production domain considering non-linear and time-varying dependencies. Through its results and identification of the current state-of-the-art, this study provides a theoretical contribution to the scientific discourse. Additionally, practitioners can utilize these findings as an indicator for suitable methods for a specific application and problem context. Consequently, this paper is situated between

SLRs focusing on the applicability of general concepts [7-9] and SLRs identifying state-of-the-art methods for specific applications [10-14] and complements these related works.

However, certain limitations regarding the research design and selection criteria are present. First, while providing a definition of generic industrial forecasting applications allows to determine general trends in the applicability of data-driven categories, the actual applicability is determined by the specifics of the application and problem domain. Therefore, the results can only act as an indicator and not as the sole criteria for selecting appropriate forecasting models for a specific application. Second, even though this study extends the considered domain to the production context in general, other applications, such as sales [63] or temperature forecast [64], are subject to non-linear and time-dependent behavior as well. Therefore, these related fields could also provide interesting methods for the industrial sector. Third, since this study is entirely focused on data-driven approaches, well-known physical models to dynamically predict future behavior via partial differential equations are not considered. However, hybrid methods, such as physics-informed neural networks, combine advantages from both research fields, hence, leading to interesting results [65]. Considering these research limitations and the quantitative findings in this work, the following future research direction can be derived:

- Since three economic sectors currently dominate the application of industrial forecasting models, other industries could adapt forecasting methods and benefit from these approaches.
- Applications for yield prediction and process optimization can profit from other applications which represent the majority of reviewed literature sources.
- Semi-supervised learning models led to competitive results, however, are currently underrepresented in industrial forecasting applications. Thus, hybrid approaches, for instance with NN based auto-encoders, are an active field of research.
- Although reinforcement learning is currently not applied for industrial forecasting models, process optimization applications could particularly benefit from this learning type in the future [8, 62].
- Even though CNN based models are originally developed for image processing, the characteristics can be successfully exploited for time series data as well. Therefore, the field of industrial forecasting can benefit from related disciplines of image and video processing/analysis.

In conclusion, forecasting in production systems is actively researched at a fast pace. Correspondingly, the findings of this SLR provide insights into this domain to foster an understanding of the current state-of-the-art for industrial forecasting to facilitate future research initiatives.

References

[1] H. M. Elhusseiny and J. Crispim, "Smes, barriers and opportunities on adopting industry 4.0: A review." Procedia Computer Science, vol. 196, pp. 864–871, 2022.

[2] G. Piscitelli, A. Ferazzoli, A. Petrillo, R. Cioffi, A. Parmentola, and M. Travaglioni, "Circular economy models in the industry 4.0 era: A review of the last decade," Procedia Manufacturing, vol. 42, pp. 227–234, 2020.

[3] J. M. M ̈uller and K.-I. Voigt, "Sustainable industrial value creation in smes: A comparison between industry 4.0 and made in china 2025," International Journal of Precision Engineering and Manufacturing-Green Technology, vol. 5, pp. 659–670, 10 2018.

[4] L. Li, "China's manufacturing locus in 2025: With a comparison of "made-in-china 2025" and "industry 4.0"," Technological Forecasting and Social Change, vol. 135, pp. 66–74, 10 2018.

[5] D. Trotta and P. Garengo, "Industry 4.0 key research topics: A bib-liometric review," in 2018 7th International Conference on Industrial Technology and Management (ICITM), vol. 2018-Janua. IEEE, 3 2018, pp. 113–117.

[6] L. Cattaneo, L. Fumagalli, M. Macchi, and E. Negri, "Clarifying data analytics concepts for industrial engineering," IFAC-PapersOnLine, vol. 51, pp. 820–825, 1 2018.

[7] M. M. Rathore, S. A. Shah, D. Shukla, E. Bentafat, and S. Bakiras, "The role of ai, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities," pp. 32 030–32 052, 2021.

[8] M. Panzer and B. Bender, "Deep reinforcement learning in production systems: a systematic literature review," International Journal of Production Research, pp. 1–26, 9 2021.

[9] Z. Kang, C. Catal, and B. Tekinerdogan, "Machine learning applications in production lines: A systematic literature review," Computers Industrial Engineering, vol. 149, p. 106773, 11 2020.

[10] T. P. Carvalho, F. A. A. M. N. Soares, R. Vita, R. da P. Francisco, J. P. Basto, and S. G. S. Alcalá, "A systematic literature review of machine learning methods applied to predictive maintenance," Computers Industrial Engineering, vol. 137, p. 106024, 11 2019.

[11] W. Zhang, D. Yang, and H. Wang, "Data-driven methods for predictive maintenance of industrial equipment: A survey," IEEE Systems Journal, vol. 13, pp. 2213–2227, 9 2019.

[12] L. F. F. Barbosa, A. Nascimento, M. H. Mathias, and J. A. de Carvalho, "Machine learning methods applied to drilling rate of penetration prediction and optimization - a review,"

Journal of Petroleum Science and Engineering, vol. 183, p. 106332, 12 2019.
[13] D. Weichert, P. Link, A. Stoll, S. Rüping, S. Ihlenfeldt, and S. Wrobel, "A review of machine learning for the optimization of production processes," The International Journal of Advanced Manufacturing Technology, vol. 104, pp. 1889–1902, 10 2019.
[14] M. Fernandes, J. M. Corchado, and G. Marreiros, "Machine learning techniques applied to mechanical fault diagnosis and fault prognosis in the context of real industrial manufacturing use-cases: a systematic literature review," Applied Intelligence, 3 2022.
[15] L. Breiman, "Statistical modeling: The two cultures (with comments and a rejoinder by the author)," Statistical Science, vol. 16, pp. 199–231, 8 2001.
[16] X. Dastile, T. Celik, and M. Potsane, "Statistical and machine learning models in credit scoring: A systematic literature survey," Applied Soft Computing, vol. 91, p. 106263, 6 2020.
[17] S. M. Cho, P. C. Austin, H. J. Ross, H. Abdel-Qadir, D. Chicco, G. Tomlinson, C. Taheri, F. Foroutan, P. R. Lawler, F. Billia, A. Gramolini, S. Epelman, B. Wang, and D. S. Lee, "Machine learning compared with conventional statistical models for predicting myocardial infarction readmission and mortality: A systematic review," Canadian Journal of Cardiology, vol. 37, pp. 1207–1214, 8 2021.
[18] B. Grillone, S. Danov, A. Sumper, J. Cipriano, and G. Mor, "A review of deterministic and data-driven methods to quantify energy efficiency savings and to predict retrofitting scenarios in buildings," Renewable and Sustainable Energy Reviews, vol. 131, p. 110027, 10 2020.
[19] X.-H. Dang, S. Y. Shah, and P. Zerfos, "seq2graph: Discovering dynamic non-linear dependencies from multivariate time series," in 2019 IEEE International Conference on Big Data (Big Data). IEEE, 12 2019, pp. 1774–1783.
[20] R. N. Thomas and R. Gupta, "A survey on machine learning approaches and its techniques:," in 2020 IEEE International Students' Conference on Electrical,Electronics and Computer Science (SCEECS). IEEE, 2 2020, pp. 1–6.
[21] G.-J. Qi and J. Luo, "Small data challenges in big data era: A survey of recent progress on unsupervised and semi-supervised methods," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8828, pp. 1–1, 2020.
[22] T. Meng, X. Jing, Z. Yan, and W. Pedrycz, "A survey on machine learning for data fusion," Information Fusion, vol. 57, pp. 115–129, 2020.
[23] R. He, X. Li, G. Chen, G. Chen, and Y. Liu, "Generative adversarial network-based semi-supervised learning for real-time risk warning of process industries," Expert Systems with Applications, vol. 150, pp. 1–12, 7 2020.
[24] N. K. Manjunath, A. Shiri, M. Hosseini, B. Prakash, N. R. Waytowich, and T. Mohsenin, "An energy efficient edgeai autoencoder accelerator for reinforcement learning," IEEE Open Journal of Circuits and Systems, vol. 2, pp. 182–195, 2021.
[25] R. Nian, J. Liu, and B. Huang, "A review on reinforcement learning: Introduction and applications in industrial process control," Computers Chemical Engineering, 4 2020.
[26] A. Pan, W. Xu, L. Wang, and H. Ren, "Additional planning with multiple objectives for reinforcement learning," Knowledge-Based Systems, vol. 193, p. 105392, 2020.
[27] Z. Ge, Z. Song, S. X. Ding, and B. Huang, "Data mining and analytics in the process industry: The role of machine learning," IEEE Access, vol. 5, pp. 20 590–20 616, 2017.
[28] S. Singh and S. Srivastava, "Review of clustering techniques in control system," Procedia Computer Science, vol. 173, pp. 272–280, 2020.
[29] X. Xu, T. Liang, J. Zhu, D. Zheng, and T. Sun, "Review of classical dimensionality reduction and sample selection methods for large-scale data processing," Neurocomputing, vol. 328, pp. 5–15, 2 2019.
[30] V. Bol´on-Canedo and A. Alonso-Betanzos, "Ensembles for feature selection: A review and future trends," Information Fusion, vol. 52, pp. 1–12, 12 2019.
[31] B. Kitchenham, "Procedures for performing systematic reviews," 2004.
[32] Y. Yang, Y. Song, Y. An, Y. Li, and H. Wang, "A general data renewal model for prediction algorithms in industrial data analytics," in 2019 IEEE International Conference on Power Data Science (ICPDS). IEEE, 11 2019, pp. 43–50.
[33] G. Wang, A. Ledwoch, R. M. Hasani, R. Grosu, and A. Brintrup, "A generative neural network model for the quality prediction of work in progress products," Applied Soft Computing, vol. 85, p. 105683, 12 2019.
[34] X. Liu, J. Jin, W. Wu, and F. Herz, "A novel support vector machine ensemble model for estimation of free lime content in cement clinkers," ISA Transactions, vol. 99, pp. 479–487, 4 2020.
[35] R. Meyes, J. Donauer, A. Schmeing, and T. Meisen, "A recurrent neural network architecture for failure prediction in deep drawing sensory time series data," Procedia Manufacturing, vol. 34, pp. 789–797, 2019.
[36] Y. Yang, Y. Bai, C. Li, and Y.-N. Yang, "Application research of arima model in wind turbine gearbox fault trend prediction," in 2018 International Conference on Sensing,Diagnostics, Prognostics, and Control (SDPC). IEEE, 8 2018, pp. 520–526.
[37] E. Lughofer, A.-C. Zavoianu, R. Pollak, M. Pratama, P. Meyer-Heye, H. Zörrer, C. Eitzinger, and T. Radauer, "Autonomous supervision and optimization of product quality in a multi-stage manufacturing process based on self-adaptive prediction models," Journal of Process Control, vol. 76, pp. 27–45, 4 2019.
[38] S. Ding, H. Yang, Z. Wang, G. Song, Y. Peng, and X. Peng, "Dynamic prediction of the silicon content in the blast furnace using lstm-rnn-based models," in 2018 International Computers, Signals and Systems Conference (ICOMSSC). IEEE, 9 2018, pp. 491–495.
[39] A. P. F. Machado, C. Z. Resende, and D. C. Cavalieri, "Estimation and prediction of motor load torque applied to electrical submersible pumps," Control Engineering Practice, vol. 84, pp. 284–296, 3 2019.
[40] E. Lughofer, R. Pollak, A.-C. Zavoianu, P. Meyer-Heye, H. Zörrer, C. Eitzinger, J. Lehner, T. Radauer, and M. Pratama, "Evolving time-series based prediction models for quality criteria in a multi-stage production process," in 2018 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS). IEEE, 5 2018, pp. 1–10.
[41] F. Hoppe, J. Hohmann, M. Knoll, C. Kubik, and P. Groche, "Feature-based supervision of shear cutting processes on the basis of force measurements: Evaluation of feature engineering and feature extraction," Procedia Manufacturing, vol. 34, pp. 847–856, 2019.
[42] Z. Huang, J. Zhu, J. Lei, X. Li, and F. Tian, "Tool wear predicting based on multisensory raw signals fusion by reshaped time series convolutional neural network in manufacturing," IEEE Access, vol. 7, pp. 178 640–178 651, 2019.
[43] A. Pasias, T. Vafeiadis, D. Ioannidis, and D. Tzovaras, "Forecasting bath and metal height features in electrolysis process," in 2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS). IEEE, 5 2019, pp. 312–317.
[44] W. Liu, W. D. Liu, and J. Gu, "Forecasting oil production using ensemble empirical model decomposition based long

short-term memory neural network," Journal of Petroleum Science and Engineering, vol. 189, p. 107013, 6 2020.

[45] C.-H. Yeh, Y.-C. Fan, and W.-C. Peng, "Interpretable multi-task learning for product quality prediction with attention mechanism," in 2019 IEEE 35th International Conference on Data Engineering (ICDE), vol. 2019-April. IEEE, 4 2019, pp. 1910–1921.

[46] M. Li, Y. Han, Y. Huo, and Q. Zhao, "Long short-term memory based on random forest-recursive feature eliminated for hot metal silcion content prediction of blast furnace," in 2019 IEEE 5th International Conference on Computer and Communications (ICCC). IEEE, 12 2020, pp. 1862–1866.

[47] J. Zhang, P. Wang, R. Yan, and R. X. Gao, "Long short-term memory for machine remaining life prediction," Journal of Manufacturing Systems, vol. 48, pp. 78–86, 7 2018.

[48] D. A. Sala, A. Jalalvand, A. V. Y.-D. Deyne, and E. Mannens, "Multivariate time series for data-driven endpoint prediction in the basic oxygen furnace," in 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, 12 2018, pp. 1419–1426.

[49] Z. Lv, J. Zhao, Y. Zhai, and W. Wang, "Non-iterative t–s fuzzy modeling with random hidden-layer structure for bfg pipeline pressure prediction," Control Engineering Practice, vol. 76, pp. 96–103, 7 2018.

[50] X. Yuan, L. Li, and Y. Wang, "Nonlinear dynamic soft sensor modeling with supervised long short-term memory network," IEEE Transactions on Industrial Informatics, vol. 16, pp. 3168–3176, 5 2020.

[51] P. Jia, H. Liu, S. Wang, and P. Wang, "Research on a mine gas concentration forecasting model based on a gru network," IEEE Access, vol. 8, pp. 38 023–38 031, 2020.

[52] J. G. C. Pena, V. B. de Oliveira, and J. L. F. Salles, "Optimal scheduling of a by-product gas supply system in the iron- and steel-making process under uncertainties," Computers Chemical Engineering, vol. 125, pp. 351–364, 6 2019.

[53] E. Lughofer, R. Pollak, A.-C. Zavoianu, M. Pratama, P. Meyer-Heye, H. Zörrer, C. Eitzinger, J. Haim, and T. Radauer, "Self-adaptive evolving forecast models with incremental pls space updating for on-line prediction of micro-fluidic chip quality," Engineering Applications of Artificial Intelligence, vol. 68, pp. 131–151, 2 2018.

[54] C.-L. Liu, W.-H. Hsaio, and Y.-C. Tu, "Time series classification with multivariate convolutional neural network," IEEE Transactions on Industrial Electronics, vol. 66, pp. 4788–4797, 6 2019.

[55] D. Moldovan, I. Anghel, T. Cioara, and I. Salomie, "Time series features extraction versus lstm for manufacturing processes performance prediction," in 2019 International Conference on Speech Technology and Human-Computer Dialogue (SpeD). IEEE, 10 2019, pp. 1–10.

[56] X. Song, Y. Liu, L. Xue, J. Wang, J. Zhang, J. Wang, L. Jiang, and Z. Cheng, "Time-series well performance prediction based on long short-term memory (lstm) neural network model," Journal of Petroleum Science and Engineering, vol. 186, p. 106682, 3 2020.

[57] A. Sagheer and M. Kotb, "Time series forecasting of petroleum production using deep lstm recurrent networks," Neurocomputing, vol. 323, pp. 203–213, 1 2019.

[58] C.-Y. Lin, Y.-M. Hsieh, F.-T. Cheng, H.-C. Huang, and M. Adnan, "Time series prediction algorithm for intelligent predictive maintenance," IEEE Robotics and Automation Letters, vol. 4, pp. 2807–2814, 7 2019.

[59] D. Zhao, K. Hao, L. Chen, B. Wei, X.-S. Tang, X. Cai, and L. Ren, "Trend-gru model based time series data prediction in melt transport process," in 2019 19th International Conference on Control, Automation and Systems (ICCAS), vol. 2019-Octob. IEEE, 10 2019, pp. 541–546.

[60] Eurostat, "Statistical classification of economic activities in the european community, rev. 2 (2008)," 2008. [Online]. Available: https://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm? TargetUrl=LST NOM DTL&StrNom=NACE REV2

[61] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," Technometrics, vol. 12, 1970.

[62] E. Pan, P. Petsagkourakis, M. Mowbray, D. Zhang, and E. A. del Rio-Chanona, "Constrained model-free reinforcement learning for process optimization," Computers Chemical Engineering, vol. 154, p. 107462, 2021.

[63] C. Zhang, Y. X. Tian, and Z. P. Fan, "Forecasting sales using online review and search engine data: A method based on pca–dsfoa–bpnn," International Journal of Forecasting, 2021.

[64] X. Meng and J. W. Taylor, "Comparing probabilistic forecasts of the daily minimum and maximum temperature," International Journal of Forecasting, vol. 38, pp. 267–281, 1 2022.

[65] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," Nature Reviews Physics, vol. 3, pp. 422–440, 6 2021.

[66] A. Şahin, A. D. Sayımlar, Z. M. Teksan, and E. Albey, "A markovian approach for time series prediction for quality control," IFAC-PapersOnLine, vol. 52, pp. 1902–1907, 2019.

[67] K.-T. Lee, Y.-S. Lee, and H. Yoon, "Development of edge-based deep learning prediction model for defect prediction in manufacturing process," in 2019 International Conference on Information and Communication Technology Convergence (ICTC). IEEE, 10 2019, pp. 248–250.

[68] H. Bisri and M. L. Singgih, "Improvement of shewhart control chart for auto-correlated data in continuous production process," in 2018 4th International Conference on Science and Technology (ICST). IEEE, 8 2018, pp. 1–6.

[69] C. Luo, C. Tan, and Y. Zheng, "Long-term prediction of time series based on stepwise linear division algorithm and time-variant zonary fuzzy information granules," International Journal of Approximate Reasoning, vol. 108, pp. 38–61, 5 2019.

[70] B. Namoano, A. Starr, C. Emmanouilidis, and R. C. Cristobal, "Online change detection techniques in time series: An overview," in 2019 IEEE International Conference on Prognostics and Health Management (ICPHM). IEEE, 6 2019, pp. 1–10.

[71] X. Zhang, L. Zhang, H. Chen, and B. Dai, "Prediction of coal feeding during sintering in a rotary kiln based on statistical learning in the phase space," ISA Transactions, vol. 83, pp. 248–260, 2018.

[72] Y. Han, C. Fan, M. Xu, Z. Geng, and Y. Zhong, "Production capacity analysis and energy saving of complex chemical processes using lstm based on attention mechanism," Applied Thermal Engineering, vol. 160, p. 114072, 9 2019.

[73] G. Sbrana and A. Silvestrini, "Random switching exponential smoothing: A new estimation approach," International Journal of Production Economics, vol. 211, pp. 211–220, 5 2019.

[74] K. Sujatha, N. Bhavani, V. Srividhya, V. Karthikeyan, and N. Jayachitra, "Soft sensor with shape descriptors for flame quality prediction based on lstm regression," pp. 115–138, 2020.

Appendix A

Table 5 contains literature sources, which were not considered relevant and excluded after the full-test assessment.

Table 5. Reasons for excluding full text assessed literature sources

Reference	Reason for exclusion
[66]	Reactive approach is not suited for forecasting scenarios
[67]	No time series data considered
[68]	Focused on optimizing control charts instead of forecasting
[69]	Not related to production processes
[70]	General review of change detection techniques without explicit application
[71]	Specifically developed for rotary kiln processes and not suited for the research questions
[72]	Focused on identifying optimization potentials in chemical processes
[73]	Not related to production processes
[74]	Focused on image data as input for predictions which is not applicable for industrial forecasting