Review

Not peer-reviewed version

Tool Condition Monitoring Under Different Operating Conditions Using ML with Scarce Data: A Review

Maryam Assafo * and Peter Langendoerfer

Posted Date: 4 May 2026

doi: 10.20944/preprints202605.0129.v1

Keywords: different operating conditions; domain adaptation; domain generalization; machine learning; meta-learning; scarce data; tool condition monitoring; transfer learning

Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 May 2026

doi:10.20944/preprints202605.0129.v1

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

Review

Tool Condition Monitoring Under Different Operating Conditions Using ML with Scarce Data: A Review

Maryam Assafo * and Peter Langendoerfer

Department of Wireless Systems, BTU Cottbus-Senftenburg, Cottbus, Germany
* Correspondence: maryam.assafo@b-tu.de

Abstract

Tool condition monitoring (TCM) is essential for ensuring good quality products, machining reliability, efficiency, and sustainability. Machine learning (ML), including deep learning (DL), has been extensively used in the literature to address different TCM tasks such as tool state recognition, tool wear prediction, and remaining useful life prediction. Nevertheless, the adoption of existing methods in real-world manufacturing is still hurdled by different practical challenges. This paper focuses on two key challenges facing ML-based TCM: 1) Variability of operating conditions; 2) Data scarcity. The first challenge arises from the variations in data distributions (domain shift) caused under cross operating conditions, which leads a trained ML model to generalize poorly on data from operating conditions unseen during training. The second challenge stems from the impracticality of collecting sufficient data on real-world shop floors, especially when labeled data is needed. Addressing simultaneously both challenges inherently leads to problem and evaluation settings different from those concerning only one single challenge without having the constraints imposed by the other (e.g., addressing the domain shift assuming the availability of sufficient data, or addressing the data scarcity under single-operating conditions). This paper presents a review focused on TCM considering both different operating conditions and data scarcity scenarios. The works reviewed are based on adaptation- and/or generalization-oriented solutions leveraging prior knowledge across various related-yet-distinct learning settings, namely transfer learning, domain adaptation, domain generalization, meta-learning, and hybrid settings. Future research opportunities are also presented. This review can serve as a guide for both researchers and practitioners, presenting state-of-the-art practices and concrete insights to tackle and advance the challenging industry application of TCM.

Keywords: different operating conditions; domain adaptation; domain generalization; machine learning; meta-learning; scarce data; tool condition monitoring; transfer learning

---

1. Introduction

Cutting tools are key industrial components in machining, used to remove materials of a workpiece to obtain products that meet predefined specifications [1]. The tool health state directly influences the product quality, machining efficiency, reliability, and sustainability [2–4], making tool condition monitoring (TCM) a critical task in modern manufacturing [1,4]. TCM methods can be categorized into direct and indirect methods, depending on the type of variables measured to monitor the tool health state. Direct methods are based on measuring tool geometry-related variables, e.g., dimension of the worn area, using special equipment such as cameras. Despite their high potential accuracy, these methods are deemed infeasible in practice since capturing and inspecting the actual tool geometry requires frequent stoppages of the machine, which negatively affects machining efficiency. Conversely, indirect methods are based on measuring physical quantities that can be indicative of the tool state, e.g., cutting forces. To this end, one or more sensors are usually deployed to acquire the corresponding signals which are subsequently processed and analyzed by data analytics techniques to finally infer the tool health state [5]. Being a powerful tool enabling automatic learning from complex data, machine

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 May 2026 doi:10.20944/preprints202605.0129.v1

learning (ML) has been increasingly used for performing different data-driven TCM tasks, including anomaly detection (e.g., tool breakage detection), tool state recognition, tool wear prediction, and remaining useful life (RUL) prediction [1]. Despite the state-of-the-art predictive performance of many ML algorithms existing in the literature, including both traditional ML and deep learning (DL), they are still narrowly adopted in real-world manufacturing. This paper focuses on two critical challenges facing ML-based TCM, namely the variable operating conditions and data scarcity.

TCM under different operating conditions: In real-world manufacturing, operating conditions (e.g., spindle rotation speed, feed rate, depth of cut, workpiece, tool, etc.) are not fixed but rather changing in accordance with production requirements and machining dynamics. In the context of ML-based TCM, changing operating conditions causes changes in the tool health degradation rate and monitoring sensor data used by ML/DL (whether raw signals or features) [6], leading to variations in the marginal and/or conditional distributions [7]. When such variations exist between training and test data (i.e., domain shift), a trained model may suffer from a poor cross-condition generalization, i.e., a degraded predictive performance when tested on data taken under operating conditions unseen during training, making the model unreliable under variable operating conditions [1]. While changes in any of the operating-condition variables creates a different operating condition, changing certain variables has particularly shown to be more influential and challenging for TCM than others. For example, it was found in [7] that the variations in data marginal distribution caused by changing the workpiece material and tool material were larger than those caused by changing cutting parameters such as feed rate and depth of cut. In addition, it was found in [8] that multi-operating-condition models trained on data samples taken under the same feed rate and depth of cut but different workpiece materials has a lower accuracy compared to models trained on data samples taken under different feed rate or depth of cut but shared workpiece material, indicating that meaningful data patterns for TCM under varying workpiece materials are more challenging to capture and model. Moreover, different operating-condition variables may simultaneously change. In summary, the difficulty of addressing the challenge imposed by changing operating conditions may vary depending on the resulting change in data.

Data scarcity: It is primarily caused by the impracticality and cost of obtaining sufficient data on real-world shop floors [7,9]. This is especially true when large, labeled data is required to accurately train the ML model, which is commonly the case with DL algorithms [6]. Moreover, it is often prohibitive to collect adequate data across all the operating conditions of interest [10]. As such, while data scarcity exists even in single-operating-condition scenarios, it is massively more pronounced in multi-operating-condition settings.

Each of the two aforementioned challenges has received a lot of interest recently. However, addressing simultaneously both challenges inherently leads to problem and evaluation settings different from those concerning only one single challenge without having the constraints imposed by the other. For example, different solutions addressing the variability of operating conditions are based on training single-operating-condition models and/or multi-operating-condition models from scratch, assuming there is sufficient data across different operating conditions, as can be seen in the review presented in [1]. On the other hand, many works that attempt to overcome the data scarcity challenge consider only a single, fixed operating condition, neglecting the cross-condition generalization aspect, e.g., [11]. Since addressing both challenges is essential for the wide adoption of ML-based TCM in real-world manufacturing, this paper presents a review concerning TCM under different operating conditions and data scarcity scenarios. Unlike the already-existing review paper [1] concerning TCM under different operating conditions using training from scratch, the works reviewed in this paper are based on leveraging prior knowledge learned from source domain/task(s) to handle data scarcity in the target domain. Different data scarcity settings are considered for training, namely: 1) Limited labeled target data; 2) Limited unlabeled target data; 3) Absence of target data; 4) Limited source data. The solutions presented are largely generalization- and/or adaptation-oriented, and are fundamentally based on unique learning paradigms and/or settings, namely transfer learning (TL) [12], domain

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

adaptation (DA) [13], domain generalization (DG) [3], meta-learning [14], and hybrid settings [7]. Figure 1 summarizes the challenges and solutions presented in this paper and Table 1 presents an overview of the works reviewed in this paper. Moreover, this paper also provides feature research opportunities to further advance this area.

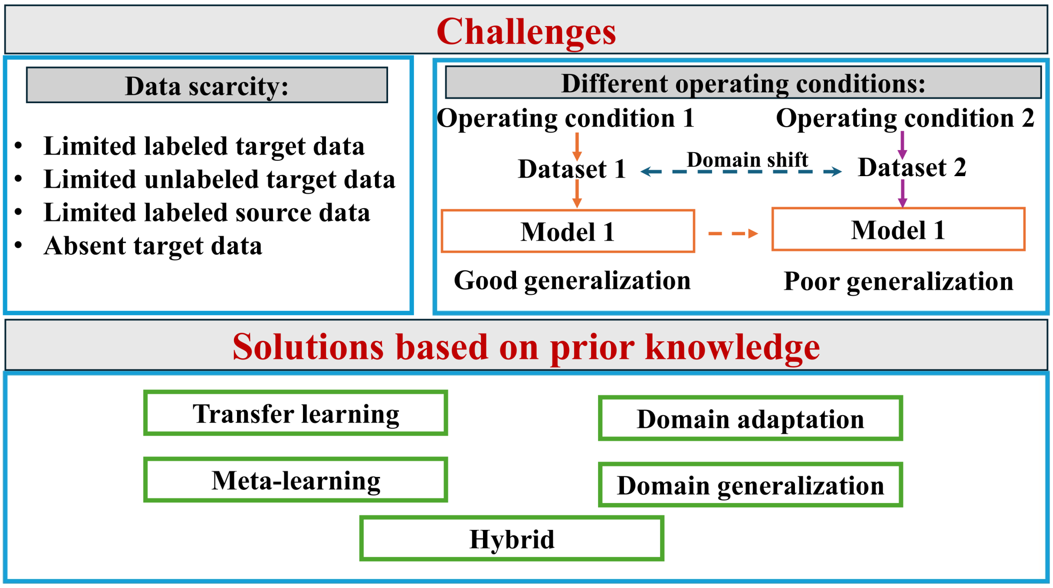

Figure 1. Challenges and solutions presented in this paper.

The remainder of this paper is as follows. Section 2 presents TCM solutions based on prior knowledge. Section 3 discusses the impact of integrating contextual information. Section 4 presents conclusions and future opportunities.

Table 1. Overview of the works reviewed in this paper.

Reference	TCM task	ML task	Learning setting
[10,12,15]	Tool state recognition	Multiclass classification	Transfer learning
[16]	Remaining useful life prediction	Regression	Transfer learning
[17]	Tool wear prediction	Regression	Transfer learning
[18], [19]	Tool state recognition	Multiclass classification	Domain adaptation
[4]	Remaining useful life prediction	Regression	Domain adaptation
[13,20–23]	Tool wear prediction	Regression	Domain adaptation
[2]	Tool state recognition	Multiclass classification	Domain generalization
[3,24]	Tool wear prediction	Regression	Domain generalization
[25]	Tool state recognition	Multiclass classification	Meta-learning
[14]	Tool wear prediction	Regression	Meta-learning
[7]	Tool wear prediction	Regression	Meta-learning & domain adaptation
[9,26,27]	Tool wear prediction	Regression	Meta-learning & domain generalization
[28]	Tool state recognition	Multiclass classification	Meta-learning & domain generalization

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 May 2026 doi:10.20944/preprints202605.0129.v1

2. TCM Based on Prior Knowledge

Traditional ML learning paradigms require a sufficient amount of data for model training, especially for DL algorithms, and they are usually based on learning a specific task [7]. The ML-based solutions presented in this section are based on learning paradigms and settings that can address the problem of data scarcity and variable operating conditions by leveraging a previously acquired knowledge (in the source domain) to enhance the performance in the target domain, and thus avoiding the need for training a model from scratch to fit the target data. Each domain corresponds to a specific operating condition. These learning paradigms and settings are: TL, DA, DG, meta-learning, and hybrid. While they may be closely related, they differ in their data requirements and ultimate objectives, offering distinct use cases. For example, both DA and DG address the problem of domain shift, yet DA focuses on specific target domain(s) while DG concerns the generalization to any unseen target domain. Additionally, unlike DA, DG has no access to target data during training [29]. Moreover, some settings require, by definition, more than one source task (e.g., meta-learning), whereas others can be applied to both single- and multisource domains (e.g., DG and DA), as will be shown in more detail below.

2.1. Transfer Learning (TL)

TL aims to transfer knowledge learned in a particular domain/task (source domain) to new, related domain/task (target domain). Compared to training from scratch, leveraging models trained on large, diverse datasets presents a vital solution for building effective TCM models, especially under limited data in the target domain [15]. Image data is frequently used as input in TL-based TCM works. To this end, different image types exist, including inspection images of the tool [17], time-frequency representations of sensor signals (e.g., scalograms generated by the continuous wavelet transform [12,15]), time series-converted images [10]. Regarding source datasets, they may be closely related to the target domain (e.g., tool monitoring images [17]), or completely different (e.g., natural images [12]). Considering the latter, CNN models previously trained on ImageNet dataset for image classification are popular, e.g., GoogLeNet, SqueezeNet, VGG, ResNet, ShuffleNet, AlexNet, DenseNet, and Inception [12,15,17]. In this case, typical TL implementations are based on retaining the original feature extraction layers of the pretrained models and only replacing the last, task-specific layers with new layers that can be trained on the target data, e.g., [12,15]. Instead of using existing, pretrained models, some works proposed TL architectures. For example, a deep TL model based on residual variation autoencoder was proposed in [10] to transfer parameter from one cutting tool to another.

The TL performance is largely influenced by the difference in feature distributions between the source and target domains, which can possibly cause negative transfer. To account for this issue, some works enable domain alignment to mitigate domain discrepancy, e.g., by adding the metric of maximum mean discrepancy (MMD) into the training loss function during fine-tuning [16,17]. Nevertheless, the choice of the source task plays a critical role in the TL performance, regardless of whether domain alignment was integrated or not, as can be seen in [16].

While TL has proven to be vital for building accurate and efficient TCM models for specific target operating conditions, the cross-condition generalization problem may still exist. For example, the performance of TCM models in [12], built using fine-tuning different pretrained CNN networks with sensory scalograms as inputs, significantly dropped under unseen operating conditions. This calls for TL-specific solutions that can yield wide generalization.

2.2. Domain Adaptation (DA)

DA is a subfield of TL that addresses the distribution difference between the source and target domains (domain shift) using labeled source data and usually unlabeled or scarcely-labeled target data [30]. Unlike TL, DA necessitates the source and target tasks to be identical [29]. Depending on the number of source domains, DA is categorized into single-source DA [18,22] and multisource DA [4,13]. Similarly, considering the number of target domains, DA is categorized into single-target DA and

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 May 2026 doi:10.20944/preprints202605.0129.v1

multitarget DA. While the majority of works consider a general DA framework, some works focus on addressing problems specific to certain settings, e.g., multisource DA [4], multitarget DA [19]. There are different approaches for DA, which can be broadly divided into instance-, feature-, and adversarial training-based [20].

Instance-based DA using TrAdaBoost regressor, which combines adaptive boosting and instance weighting, was used in [22] for tool wear prediction. The weighting mainly aims to lessen the impact of source instances that are likely to cause a negative transfer, while reinforcing target instances that can enhance the target task on unseen data [22]. Six different base estimators were tested, such as linear regression and k-nearest neighbor (KNN), among which random forest regressor showed the best performance. The role of feature selection in DA was also showcased, where different FS techniques, namely, sequential feature selection (SFS), recursive feature elimination, and recursive feature addition, were compared and the best feature selection method was eventually selected (here SFS). TrAdaBoost was also used in [4] to address multisource DA for RUL prediction. To account for the negative impact of some interactions among the source domains, a source selection strategy was implemented prior to TrAdaBoost with the aim of retaining a subset of the source domains with the highest similarity with the target domain, where the similarity between the target domain and each of the source domains was measured using dynamic time warping (DTW) and Wasserstein distance (WD). The above two instance-based approaches yield a shallow DA [29], which can be particularly useful when maintaining computational efficiency is key for the application.

Feature-based approaches are commonly used for DA, which aim to align the features between source and target domains [29]. In [13], MMD was used to align high-level deep spatial and temporal features between the source and target domains for tool wear prediction, where the MMD loss was incorporated into the overall training loss. In [18], deep coral, an unsupervised feature-based DA method which aligns the second-order statistics of the source and target feature representations, was integrated into deep extreme learning machine learning (DELM) for tool state classification. To further improve the DA performance, a central loss was added as a regularization term to the loss function of DELM to improve intra-class compactness. In [23], a backbone TL architecture was proposed based on one-dimensional convolutional neural network (1D-CNN) and transformer. The model was pretrained on the source data and then adapted to the target data by only fine-tuning the last, task-specific layers. Feature alignment across different operating conditions was achieved by adding a conditional layer normalization before the final fully-connected regression head, with the aim of yielding adaptive, condition-based standardization parameters, namely scaling and bias. Normalization parameters were determined by passing operating-condition variables through two fully-connected layers. The work in [20] addresses two common limitations of DA methods, which are especially critical in regression tasks: 1) Aligning only marginal distributions, without aligning the conditional distributions of the source and target domains. 2) Scale shift of the latent representation. To this end, the authors proposed an approach based on Inverse Gramian subspace matching (IGAM) to align latent representations of the source and target domains while considering conditional distributions.

Adversarial learning has been widely integrated into deep architectures to extract domain-invariant features [29]. In [21], domain adversarial neural network (DANN) was used for tool wear prediction. DANN is a deep architecture comprising a feature extractor, domain classifier, and label predictor. It tackles the domain shift by learning domain-invariant features, specifically by minimizing the source task loss while maximizing the domain classification loss, i.e., min–max adversarial optimization. Single- and multisource tool wear datasets were tested. Among other findings, the results showed that a better single-source DA was attained when the source dataset has higher similarity with the target dataset in terms of operating conditions. Further, for two-source DA, pretraining the model on one source dataset and fine-tuning it on the other which has high similarity with the target dataset led to a higher performance than simultaneously training the model on both datasets. In [19], a multitarget DA framework based on operating-condition variation decoupling domain adversarial DA network was proposed for tool state classification. A discriminative, generalized feature representation

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 May 2026

doi:10.20944/preprints202605.0129.v1

was achieved by a domain-shared wavelet convolutional feature extractor and a class-level prototype discrepancy alignment mechanism. The domain shift was addressed considering both source-target pairs and target-target pairs. The proposed framework outperformed several methods, e.g., DANN.

2.3. Domain Generalization (DG)

DG aims to develop models with high generalization to unseen target domains without any access to the target data (whether labels or observations) during training. The later aspect makes DG more challenging than other learning settings that have certain access to the target data during training (e.g., DA and TL) [9], but this also makes DG particularly outstanding for highly data-constrained, dynamic environments, especially under single-source DG (SSDG) implementations [24]. Existing SSDG approaches are primarily based on expanding the source data coverage [24]. In [24], SSDG based on generated feature generalization was proposed for tool wear prediction using an architecture comprising a feature extractor, a generator, and a predictor. Generalization to unseen operating conditions was particularly enabled by using the three following regularization terms in the generator loss function: MMD, mutual information, and domain discrepancy loss, which ensure that the generated samples have certain similarity to the source domain while also exhibiting controlled domain variations. Instead of fully relying on data-driven approaches, some works integrate physical information and constraints into the DG framework to enhance learning generalized features across different domains [2,3]. In [2], multisource DG was proposed for tool wear recognition where physics-informed data augmentation was used to address the imbalance class problem, rotation speed-based scaling was applied to align samples across different speeds, and physics-guided constraints were applied to learn global and local features. In [3], multisource DG approach based on a physical property-guided attention mechanism, adversarial training, and MMD was proposed for tool wear prediction.

2.4. Meta-Learning (Learning to Learn)

It is a learning technique that trains a model on multiple tasks so that it develops the ability to rapidly and accurately adapt to new tasks with only small training samples. Compared to other learning paradigms such as TL, meta-learning requires fewer target data samples for adaptation, making it prominent for few-shot learning [7]. This is particularly useful for highly-dynamic manufacturing settings, e.g., small-batch and multi-variety production. Considering TCM under different operating conditions, TCM under a particular operating condition is considered a task, and thus a model is trained on different TCM tasks corresponding to different operating conditions so that it can quicky learn to perform TCM under a new operating condition. Meta-learning can be categorized into model-based, metric-based, and optimization-based, among which the latter is considered the most popular, which is based on learning shared initialization parameters among tasks [28].

In [14], a meta-learning approach using model agnostic meta-learning (MAML) was proposed for tool wear prediction. Sensory features were first extracted and only a feature subset was then selected using an FS scheme. Fully-connected neural networks (FCNN) were adopted to build base models. Different numbers of target training samples were evaluated, namely 1, 3, and 5. The results showed that good adaptation results were obtained with only one sample. In [25], an adaptive meta-level gradient update mechanism was proposed to dynamically adjust gradient weights based on the multiple tasks, where contextual data and sensor signals were used as inputs.

2.5. Hybrid Settings

Some works consider combined learning settings to leverage the advantages and/or overcome the limitations of individual settings. For example, in [7], DA (based on invariant feature space) was combined with meta-learning to address the limitation of meta-learning when large operating-condition variations (e.g., different cutting tools) exist between the source and target datasets. In [26] and [9], meta-learning was combined with multisource DG for tool wear prediction to enhance the performance of DG under scarce data in source domains. A model-agnostic domain generalization

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

(MADG) based on meta-learning and DG was proposed in [27]. In [28], a novel meta-learning approach based on similarity-based meta-representation learning was proposed for DG concerning TCM in ultrasonic metal welding.

3. Impact of Integrating Contextual Information

Contextual information representing the operating conditions, e.g., feed rate, has been widely used across different learning settings addressing data scarcity, such as meta-learning [14,25], DG [2]. To this end, contextual information mainly acts as input features of the ML model or input parameters of some processes, e.g., feature selection [14].

Contextual information as input features, in which case contextual information is fused with sensory signals/features at some point of the pipeline before being fed to the predictive ML model. The utility of using operating-condition variables as input features was highlighted in [26], where fusing operating variables and temporal sensory features has shown to enhance the meta-learning tool wear prediction compared to using temporal features alone. In [16], the operating variables of cutting speed, feed rate, cutting depth, and material removal rate were concatenated with CNN-extracted image features for RNN-based tool RUL prediction. The impact of removing each of these variables on the performance was evaluated. It was found that the cutting speed was the top contributor to a higher accuracy. Unlike [16,26], the operating variables in the framework proposed in [13,25] were processed before being fused with monitoring data, where raw sensor signals and operating variables underwent separate feature extraction, then the individual outputs were concatenated and fed into a unified feature extraction. In [13], the performance was compared with the case when sensor signals were used alone. Integrating contextual information has shown to improve the adaptation of the subsequently built model to a new dataset under unseen workpiece material.

Inputs of some processes in the ML pipeline, with the aim of yielding condition-informed processes of data, e.g., normalization [2,23], feature selection [14], data augmentation [2], and data split [9]. For example, in [14], feature selection based on an entropy weight-grey correlation analysis was performed for tool wear prediction under varying operating conditions. The correlation of extracted features with the tool wear and cutting conditions was evaluated. A higher importance was given to the features that are more sensitive to tool wear and less sensitive to cutting conditions.

4. Conclusions and Future Opportunities

This paper presents a review on tool condition monitoring under different operating conditions using machine learning with scarce data. Compared to considering only one challenge (i.e., either data scarcity or variability of operating conditions), addressing both challenges jointly leads inherently into unique problem settings and solutions. The solutions reviewed in this paper were primarily based on leveraging previous knowledge in the source domain(s), including TL, DA, DG, meta-learning, and hybrid settings. The most challenging data scarcity scenario seems to be when target data is completely absent in the offline phase (i.e., the case addressed by domain generalization approaches) while training source data is also limited. To this end, there is a notable trend in addressing DG using meta-learning as an underlying learning paradigm, allowing for few-shot domain generalization. Further, incorporating contextual information and using physics-informed ML are also gaining interest across the different learning settings, as they can considerably simplify extracting meaningful patterns from data across operating conditions, compensate for the lack of relevant data, and enhance interpretability and understandability. Despite the advances achieved in this challenging research area, there are still many future research opportunities. We present some as follow.

One-class classification (OCC): OCC represents a critical TCM task due to the relative simplicity of acquiring normal data and the cost of obtaining abnormal data on the shop floor. While there are few works on building OCC models from scratch for TCM under different operating conditions, e.g., [6], the literature still lacks works that leverage prior knowledge for OCC, e.g., DA and DG. Due to the presence of only one class (typically the healthy state) and the absence of other classes (typically faulty

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

states) in both source and target domains, existing solutions may fail to address OCC, which requires research work for this particular task.

Multimodal DA and DG: Existing works focus on data inputs of the same type. However, heterogeneous data is increasingly becoming an integral part of intelligent TCM systems [31]. Hence, there is a need for multidomain DA and DG solutions that can be applied on multimodel inputs, e.g., numerical sensory features and images (e.g., scalograms).

Multitarget DA: Existing DA approaches are predominantly based on single target domains, thus the DA framework should be performed for every target operating condition. Alternatively, employing multitarget DA can enhance the computational efficiency by optimizing the task for more than one target domain at once. A recent work addressing multitarget DA can be found in [19].

Federated learning (FL): FL is a distributed, collaborative learning paradigm that allows different local entities to develop a shared model by only communicating local models trained on local data without having to share the raw local data. It represents a vital solution for data scarcity and data-sharing concerns [32]. FL can be adopted for TCM under different operating conditions, where datasets from different companies can be used without sharing raw data [1]. Regarding the domain shift problem, existing solutions, e.g., for TL and DA, may become irrelevant under FL settings due to the inaccessible raw data, which calls for novel solutions. Relevant insights can be seen in [32].

Impact of sensor fusion under data scarcity: The impact of sensor fusion has been already investigated with models trained from scratch under different operating conditions [1]. However, it is yet to be explored in data scarcity scenarios with solutions based on prior knowledge such as DA and DG, by comparing the results with single-sensor deployments.

Funding: This work has been funded by the Federal Ministry for Digital and Transport (BMDV) of Germany within the “D4M” project under Grant 01FV2064A.

References

1. M. Assafo and P. Langendoerfer, “ML-Based Tool Condition Monitoring Under Different Operating Conditions Using Training From Scratch: A Review,” 2026, DOI: 10.13140/RG.2.2.23502.22082.
2. K. Ding et al., “A mechanism-guided dual-trend sample augmentation domain generalization method for wear stage identification of end mills under varying process conditions,” *Journal of Manufacturing Processes*, vol. 165, pp. 306–330, 2026.
3. T. Jiang et al., “PMSA-DGNet: A Physical Property-Guided Multiscale Spatial Attention Domain Generalization Network for Tool Wear Prediction,” *IEEE Transactions on Instrumentation and Measurement*, vol. 75, pp. 1–14, 2026.
4. B. Qiang et al., “Multi-source online transfer learning based on hybrid physics-data model for cross-condition tool health monitoring,” *Journal of Manufacturing Systems*, vol. 77, pp. 1–17, 2024.
5. K. Zhu, Y. San Wong, and G. S. Hong, “Wavelet analysis of sensor signals for tool condition monitoring: A review and some new results,” *International Journal of Machine Tools and Manufacture*, vol. 49, no. 7-8, pp. 537–553, 2009.
6. M. Assafo and P. Langendörfer, “A TOPSIS-Assisted Feature Selection Scheme and SOM-Based Anomaly Detection for Milling Tools Under Different Operating Conditions,” *IEEE Access*, vol. 9, pp. 90011–90028, 2021.
7. C. Liu et al., “A Meta-Invariant Feature Space Method for Accurate Tool Wear Prediction Under Cross Conditions,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 2, pp. 922–931, 2022.
8. M. Abubakr et al., “Sensors selection for tool failure detection during machining processes: a simple accurate classification model,” *CIRP Journal of Manufacturing Science and Technology*, vol. 32, pp. 108–119, 2021.
9. D. Wang et al., “Meta domain generalization for smart manufacturing: Tool wear prediction with small data,” *Journal of Manufacturing Systems*, vol. 62, pp. 441–449, 2022.
10. J. Ou et al., “Deep transfer residual variational autoencoder with multi-sensors fusion for tool condition monitoring in impeller machining,” *Measurement*, vol. 204, p. 112028, 2022.
11. Y. Jiang, B. Drescher, and G. Yuan, “A GAN-Based Multi-Sensor Data Augmentation Technique for CNC Machine Tool Wear Prediction,” *IEEE Access*, vol. 11, pp. 95782–95795, 2023.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 May 2026 doi:10.20944/preprints202605.0129.v1

12. A. Abdeltawab, Z. Xi, and Z. Longjia, "Dual-method approach of tool condition monitoring using sensor-based deep learning with vision-based image processing," *Journal of Manufacturing Processes*, vol. 150, pp. 224–256, 2025.
13. W. Yu et al., "Multi-condition milling cutter wear prediction based on split-channel information re-fusion and domain adaptation," *Expert Systems with Applications*, p. 128888, 2025.
14. Y. Li et al., "A novel method for accurately monitoring and predicting tool wear under varying cutting conditions based on meta-learning," *CIRP annals*, vol. 68, no. 1, pp. 487–490, 2019.
15. Y. Li et al., "A novel approach for tool condition monitoring based on transfer learning of deep neural networks using time–frequency images," *Journal of Intelligent Manufacturing*, vol. 35, no. 3, pp. 1159–1171, 2024.
16. M. Marei and W. Li, "Cutting tool prognostics enabled by hybrid CNN-LSTM with transfer learning," *The International Journal of Advanced Manufacturing Technology*, vol. 118, no. 3, pp. 817–836, 2022.
17. M. Marei, S. El Zaatari, and W. Li, "Transfer learning enabled convolutional neural networks for estimating health state of cutting tools," *Robotics and Computer-Integrated Manufacturing*, vol. 71, p. 102145, 2021.
18. Y. Wang et al., "Deep transfer learning for tool condition monitoring under different processing conditions," *The International Journal of Advanced Manufacturing Technology*, vol. 133, no. 1, pp. 507–519, 2024.
19. S. Yao et al., "Machining condition variation decoupling adversarial domain adaptation network: A multi-target domain tool wear state recognition method," *Mechanical Systems and Signal Processing*, vol. 252, p. 114293, 2026.
20. G. Kim et al., "Robust tool wear prediction under novel operating conditions via physics-guided unsupervised domain adaptation," *Advanced Engineering Informatics*, vol. 69, p. 103883, 2026.
21. Y. J. Chuang, M. C. Lu, and K. M. Li, "Impact of feature engineering and domain adaptation on tool wear prediction accuracy under variable cutting conditions," *Manufacturing Letters*, vol. 44, pp. 1232–1241, 2025.
22. V. Warke et al., "Robust tool wear prediction using multi-sensor fusion and time-domain features for the milling process using instance-based domain adaptation," *Knowledge-Based Systems*, vol. 288, p. 111454, 2024.
23. N. Zhang et al., "Condition-adaptive 1DCNN-transformer for multi-condition tool wear prediction," *The International Journal of Advanced Manufacturing Technology*, pp. 1–20, 2026.
24. Z. Li et al., "Single-source domain tool wear prediction based on generated feature generalization," *Egyptian Informatics Journal*, vol. 33, p. 100907, 2026.
25. B. Zhang et al., "Meta-learning-based approach for tool condition monitoring in multi-condition small sample scenarios," *Mechanical Systems and Signal Processing*, vol. 216, p. 111444, 2024.
26. X. Mo et al., "A meta-learning method for smart manufacturing: Tool wear prediction using hybrid information under various operating conditions," *Robotics and Computer-Integrated Manufacturing*, vol. 91, p. 102846, 2025.
27. X. Mo et al., "Few-Shot Learning for Smart Manufacturing: Tool Wear Prediction Using Out-of-Domain Data Based on Meta-Learning," *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)*, Auckland, New Zealand, 2023, pp. 1–6.
28. Y. Meng et al., "Meta-Learning-Based Domain Generalization for Cost-Effective Tool Condition Monitoring in Ultrasonic Metal Welding," *IEEE Transactions on Industrial Informatics*, vol. 21, no. 1, pp. 653–662, Jan. 2025.
29. A. Farahani et al., "A brief review of domain adaptation," *Advances in data science and information engineering: proceedings from ICDATA 2020 and IKE*, 2020, pp. 877–894.
30. S. Sun, H. Shi, and Y. Wu, "A survey of multi-source domain adaptation," *Information Fusion*, 24, pp. 84–92, 2015.
31. K. Zhu, "Big data oriented smart tool condition monitoring system," in Smart machining systems: Modelling, monitoring and informatics (pp. 361–381). Cham: Springer International Publishing, 2021.
32. X. Li et al., "Federated transfer learning in fault diagnosis under data privacy with target self-adaptation," *Journal of Manufacturing Systems*, vol. 68, pp. 523–535, 2023.

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

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.