Title: Improving Virtual Metrology Predictions via Parameter-Based Transfer Learning and Active Learning

Abstract

Unlike traditional metrology in semiconductor manufacturing, it uses physical methods to measure wafers that are both resource-intensive and time-consuming, increasing possibilities of causing defects to production of wafers. Virtual metrology (VM) predicts wafer measurements using sensor data, enabling real-time, non-intrusive monitoring of process performance in semiconductor manufacturing. In this study, we utilize sensor data collected from a Seagate manufacturing facility to validate our approach. Our study introduces a advanced approach to VM by combining regression modeling with transfer learning to enhance model generalization under varying manufacturing conditions. The framework proposed Virtual Metrology Transfer and Active Learning-driven Adaptation (VM-TALA) consists of two stages: first we fine-tuning a base model using a limited amount of labeled data from the target domain, then followed with iterative refinement via active learning (AL) in which the most uncertain predictions are identified and incorporated into the training set. This method improves prediction in the target domain, especially in cases where standalone models do not perform well. Experimental results demonstrate proposed framework significantly outperforms models trained solely on target domain data. The significant improvement of refined model achieves 62.90% in Root Mean Squared Error (RMSE) and 78.92% in Mean Absolute Error (MAE) across all evaluated contexts. AL helps to select the most appropriate sample for labelling to reduce the need for extensive datasets. The proposed method is advantageous in high-mix, low-volume (HMLV) manufacturing industry settings, where some stage or products are produced in lesser amounts. This innovative approach to VM aims to streamline semiconductor manufacturing, minimize defects, and optimize resource utilization by delivering strong, adaptable predictive capabilities.

Keywords: Transfer Learning, Domain adaptation, soft sensors, virtual metrology

1.Introduction

The rapid scaling of semiconductor manufacturing, with production volumes reaching thousands of wafers per day, demands highly automated, cost-effective, and responsive process control systems. Central to this challenge is the metrology infrastructure, which ensures wafer quality through detailed inspection across multiple processing stages. Traditional physical metrology, while accurate, is increasingly unsustainable at such volumes due to prohibitive costs, cycle time penalties, and inspection bottlenecks. Multiple metrology steps are required for the inspection of finished products in semiconductor manufacturing (Maitra, V., et al. 2024).

VM is the process of predicting quality of product without direct measurement, enabling faster decision-making and reduce the need for exhaustive physical measurement.VM is the process of predicting quality of product without direct measurement, enabling faster decision-making and reduce the need for exhaustive physical measurement. The proliferation of Industry 4.0 has established Virtual Metrology as an indispensable technology for intelligent manufacturing, particularly within the exacting domain of semiconductor fabrication. Industry 4.0 is characterized by the convergence of advanced analytics, cyber-physical systems, cloud computing, and the Industrial Internet of Things (IIoT), enabling real-time monitoring, predictive maintenance, and data-driven optimization of production processes (Bai et al., 2020; Soori, M., et al. 2023). These capabilities are transforming manufacturing operations by enhancing flexibility, improving product quality, and reducing downtime and waste. This makes VM an ideal fit for high-precision, high-throughput environments like semiconductor fabs. VM involves the estimation of a product's critical quality characteristics, such as line width, hole diameter and film thickness, by leveraging high dimensional sensor data collected in situ during the manufacturing process (Shim & Kang, 2022). This approach circumvents the significant costs, throughput limitations, and extended cycle-time associated with physical metrology, enabling real-time, wafer-to-wafer quality monitoring and advanced process control.

The efficacy of data-driven VM is fundamentally undermined by two persistent and intertwined challenges which are process drift and data scarcity (Guha et al.,2023). In high-mix, low-volume (HMLV) manufacturing, frequent changes in recipes, tooling, and equipment conditions cause domain shifts, making static VM models quickly outdated and less accurate. At the same time, collecting new labeled data is expensive and impractical, especially as new products are introduced regularly. This creates a cycle where models must adapt to new domains with limited data, but retraining robust models for each domain is not feasible. The key research challenge is to develop VM models that can quickly and accurately adapt to changing processes while minimizing the need for costly labeled data.

Current research has addressed these challenges through two primaries, yet largely isolated, avenues. TL has been employed to mitigate domain shift by adapting a model pre-trained on a data-rich source domain to a data-scarce target domain. Concurrently, AL has been utilized to combat data scarcity by intelligently querying the most informative unlabeled samples for annotation, thereby maximizing model performance for a given labeling budget (Sharath et al.,2024). However, a significant research gap exists at the intersection of these paradigms in VM. TL methodologies typically presume the availability of a small, labeled target dataset without specifying how to acquire it efficiently, while AL strategies, though data-efficient, are not inherently designed for domain adaptation. This gap is further deepened at a methodological level. The true frontier lies not merely in combining TL and AL, but in synergizing their most advanced and mutually enabling sub-disciplines.

Parameter-based or parameter-efficient fine-tuning (PEFT) which adapts only a small subset of parameters has been shown to be more effective and less prone to overfitting in few-shot scenarios than conventional full-model fine-tuning. However, its potential remains largely unexplored in this context (Wang et al.,2025) . The efficiency of such methods, which can adapt to a model with as few as five to ten samples, fundamentally alters the cost-benefit analysis of AL, making it a highly feasible and impactful strategy. This symbiotic relationship, where the extreme data efficiency of parameter-based TL unlocks the full potential of uncertainty-driven AL, has not systematically investigated VM as current approaches in virtual metrology lack an integrated framework that simultaneously incorporates a principled uncertainty-based AL module with a parameter-efficient adaptation mechanism. While recent advances in AL have demonstrated the effectiveness of uncertainty estimation in prioritizing informative samples for annotation, few systems align this with lightweight, transferable learning architectures capable of adapting to rapidly evolving process domains.

To address the gaps, this paper proposes a novel, integrated framework that synergistically combines parameter-based transfer learning and uncertainty-driven active learning to create a highly accurate, data-efficient, and adaptive VM system. The central hypothesis is that by creating a closed loop where an uncertainty-driven AL engine strategically selects a minimal set of high-value wafers for labeling, which are then used by a parameter-based TL module for rapid and robust model adaptation, it is possible to achieve superior predictive performance in dynamic manufacturing environments with drastically reduced metrology overhead.

The key contributions of this work are:

- The design and implementation of a two-stage framework that first leverages TL for rapid, data-efficient initial model deployment and subsequently uses AL for continuous, cost-effective model refinement and adaptation.
- Integration of Random Forest Regression with transfer learning is showing how an ensemble-based non-parametric model can be adapted across domains by reusing tree structures and fine-tuning on limited target labels.
- Show improvement in both regression accuracy for predicting wafer quality parameters and classification performance for detecting abnormal wafers, thereby directly addressing key industrial objectives of yield enhancement and cost reduction.

In this paper, the performance of different transfer learning tasks including cross factory, cross stage, and cross measurement is evaluated with a focus on the key parameters of successful strategies. The data used for our numerical experiments was obtained from Seagate's industrial semiconductor wafer manufacturing operations (Kaitwanidvilai et al., 2025).

2.Related Work

2.1 Virtual Metrology in Manufacturing

Virtual Metrology has emerged as a cost-effective and time-efficient alternative to physical metrology in advanced manufacturing. By leveraging sensor signals, process parameters, and machine learning models, VM predicts product quality without direct physical measurements. Recent studies have shown that VM significantly reduces downtime and increases throughput, especially in semiconductor fabrication and precision machining environments. However, VM's predictive accuracy heavily depends on data quality and the generalizability of the chosen model across different production scenarios.

2.2 Transfer learning and active learning

Transfer learning has become an increasingly prominent approach in manufacturing research due to its ability to mitigate the challenges associated with scarce labeled data and variations between source and target domains. By reusing knowledge from prior models and adapting it to new tasks, TL offers a pathway to reduce training time, enhance predictive capability, and maintain model relevance under changing process conditions.

Transfer learning methodologies are broadly categorized based on how knowledge is transferred between domains which included instance-based, feature-based, or parameter-based. Instance-based transfer learning focuses on reusing or reweighting data samples from the source domain to enhance learning in the target domain. Feature-based transfer learning methods aim to learn a common feature representation that bridges the gap between source and target domains. Such approaches transform or map data from both domains into a shared feature space where their distributions are more aligned, thereby enabling effective knowledge transfer. Parameter-based transfer learning involves transferring model parameters, such as weights, layers, or hyperparameters, from a source domain model to a target domain model. The pre-trained source model serves as an initial starting point, and its parameters are subsequently fine-tuned using data from the target domain.

TL applications evolved toward hybrid and physics-informed strategies. (Semitela et al.,2024) explored TL for surface defect detection in manufacturing inspection, achieving higher classification accuracy by transferring features from pre-trained vision models to sensor-based quality control tasks. Liang et al.,2025 combined convolutional and recurrent neural architectures with incremental TL for fault diagnosis in machining, enabling seamless adaptation to shift in operating conditions. (Zhu et al.,2025) integrated TL with physics-informed neural networks (TLE-PINN) for selective laser melting, allowing physical process constraints to guide melt pool morphology predictions.

A notable recent contribution is the Generative-FewShot-Active Virtual Metrology (GFA-VM) framework. This framework unifies large-scale generative modeling with few-shot fine-tuning and uncertainty-driven active sampling for adaptive process control in semiconductor manufacturing (Lin et al., 2025). It demonstrates the ability to recalibrate models with only 1-5 critical labeled samples, underscoring the extreme data efficiency achievable through such integrated approaches. A critical consideration for effective PTL, as highlighted in the literature, is the necessity for congruence between the source and target domains. The industrial processes in both domains must exhibit similar features or tasks to ensure effective knowledge transfer. This implies that while TL is powerful, its effectiveness can be limited if the source and target domains are too dissimilar, potentially leading to negative transfer. This highlights the need for careful source selection or advanced domain adaptation for disparate manufacturing processes.

AL helps reduce labeling costs by allowing models to choose which data points should be labeled next. It identifies examples that are expected to contribute the most to learning, which improves model accuracy using fewer annotations. The most common strategy is uncertainty sampling, where the model queries instances it is least confident about. This includes techniques such as entropy scoring, margin-based selection, and least-confidence sampling. Another widely used method is Query-by-Committee (QBC), which relies on prediction disagreement across multiple models to identify informative samples (Shoghi et al.,2024) A third strategy, known as expected model change, selects data points likely to have the largest impact on the model if labeled. These methods have shown strong results in real-world tasks. When combined with transfer learning, AL can refine pre-trained models with minimal additional data.

2.3 Problem Formulation: Transfer and Active Learning for Virtual Metrology

Let a source domain be denoted as $D_s \;\; = \{X_{s,} P(X_s)\}$where $X_s = \{x_1, x_2, x_3, x_4, \dots, x_m\}$

is the feature space and $P(X_s)$ is the marginal probability distribution of these features.

The source learning task $T_s = \{y_{s,} f_s(.)\}$ where y_s is the label space and $f_s(.)$ is the predictive function mapping features to labels.

For a given sample i :, $Y_s \; = fs\left(x_1(i)_{,} x_2(i), x_3(i)_{,4}^{\,x}(i), \dots, x_{4m}(i)\right) + \in$

where $\in$ represents noise or modeling error.

and a target domain

$$D_T = \{X_{s,} P(X_T)\}$$

with its target learning task

$$T_T = \{y_{s,} f_T(.)\}.$$

The objective of transfer learning is to improve the learning of the target predictive function $f_T(.)$ in D_T by exploiting knowledge from D_s and Ts , where $Ds \neq D_T$.

A domain $D = \{X, P(X)\}$ consists of a feature space X (e.g., process sensor readings) and a marginal probability distribution $P(X)$ over instances $x_i \in X$ The condition $D_s \neq D_T$ holds if either $X_s \neq X_T$ or $P(X_s) \neq P(X_T)$.

From the Virtual Metrology perspective, a task $T = \{y, f(.)\}$ consists of a label space y and a predictive function$f(.)$ modeling the conditional distribution $P(\, Y \mid X\,)$.

2.4 Combining Transfer Learning and Active Learning

The integration of transfer learning and active learning is emerging as a promising approach in manufacturing, offering both rapid model initialization and efficient refinement with minimal labeled data. TL can leverage prior models to provide a strong predictive baseline, as demonstrated in studies employing domain adaptation and transfer boosting to address domain discrepancies in industrial robotics and manufacturing processes (Ye et al., 2023). AL used selectively queries the most informative samples, thereby reducing measurement costs an especially critical factor in virtual metrology, where data acquisition can be expensive and time-consuming. Recent work on ML-based VM pipelines has shown the feasibility of integrating TL for high-dimensional sensor data with limited measurements (Guha et al., 2023). Despite these advances, the combined application of TL and AL for VM prediction remains largely unexplored, representing a significant research gap in adapting predictive models to dynamic manufacturing environments with frequent process drifts.

3. Materials and methods

3.1. Base Model: Random Forest Regression

In this work, the predictive function $f(.)$ from Section 2 is instantiated as a Random Forest Regressor. Given a feature vector $x = [x_1, x_2, \ldots, x_n]^\top$, the Random Forest model $f_{RF}(\cdot)$ consists of M individual regression trees $\{h_m(\cdot)\}_{m=1}^{M}$. Each tree is trained on a bootstrap sample of the source domain data D_S , and at each split, a random subset of features is considered.

The prediction for an input X is given by:

$$\hat{y} = f_{RF}(x) = \frac{1}{M}\sum_{m=1}^{M} h_m(x)$$

where $h_m(x)$ is the prediction of the $m-th$ regression tree.

3.2. Transfer Learning Implementation

The proposed strategy operates in two principal phases which are pre-training and fine-tuning. In the first phase, a Random Forest model is trained using extensive, high-quality historical measurements from the source domain. This model, denoted as f_s , consists of an ensemble of B decision trees, each producing a separate prediction. The overall model output is obtained as:

where $f_s(x) = \frac{1}{B}\sum_{b=1}^{B} T_b\,(x)$,

is the prediction from the $T_b(x)$ decision tree, and B While f_s achieves high accuracy in the source domain, direct application to the target domain D_T. The fine-tuning process follows an active learning paradigm. The source-trained model is applied to the target domain to produce preliminary predictions. Samples exhibiting high prediction uncertainty are identified using variance-based measures across trees in the ensemble. This active selection of informative samples accelerates domain adaptation while minimizing labeling costs.

Following the formulation in Section 2, the goal is to improve $f_T(\cdot)$ in the target domain D_T using knowledge from $f_s(\cdot)$ trained in the source domain D_s . In the Random Forest setting, this is done by:

1. Training $f_s(.)$ on $D_s = \{X_{s,}Y_s\}$
2. Initializing $f_T(\cdot)$ with the structure and learned splits from $f_s(\cdot)$.
3. Fine-tuning tree parameters with the limited labeled target domain data $D_T labeled$.

Mathematically: $f_T(\cdot) \leftarrow Adapt\left(f_s(\cdot), D_T^l abeled\right)$

3.3. Active Learning–Based Refinement

To enhance the performance of $f_T(\cdot)$ when Y_T is scarce, we employ uncertainty sampling. The variance of predictions across all trees in the ensemble is computed:

$$u(x) = \frac{1}{M} \sum_{m=1}^{M} \left(h_m(x) - \bar{h}(x)\right)^2$$

Samples with the largest $u(x)$ are selected for labeling and added to the training set. The refined model is then: $f_T^{\text{refined}}(\cdot) = \text{Retrain}\left(f_T(\cdot), D_T^{\text{labeled}} \cup D_T^{\text{new}}\right) \dashv)$ where D_T^{new} are the newly labeled high-uncertainty samples.

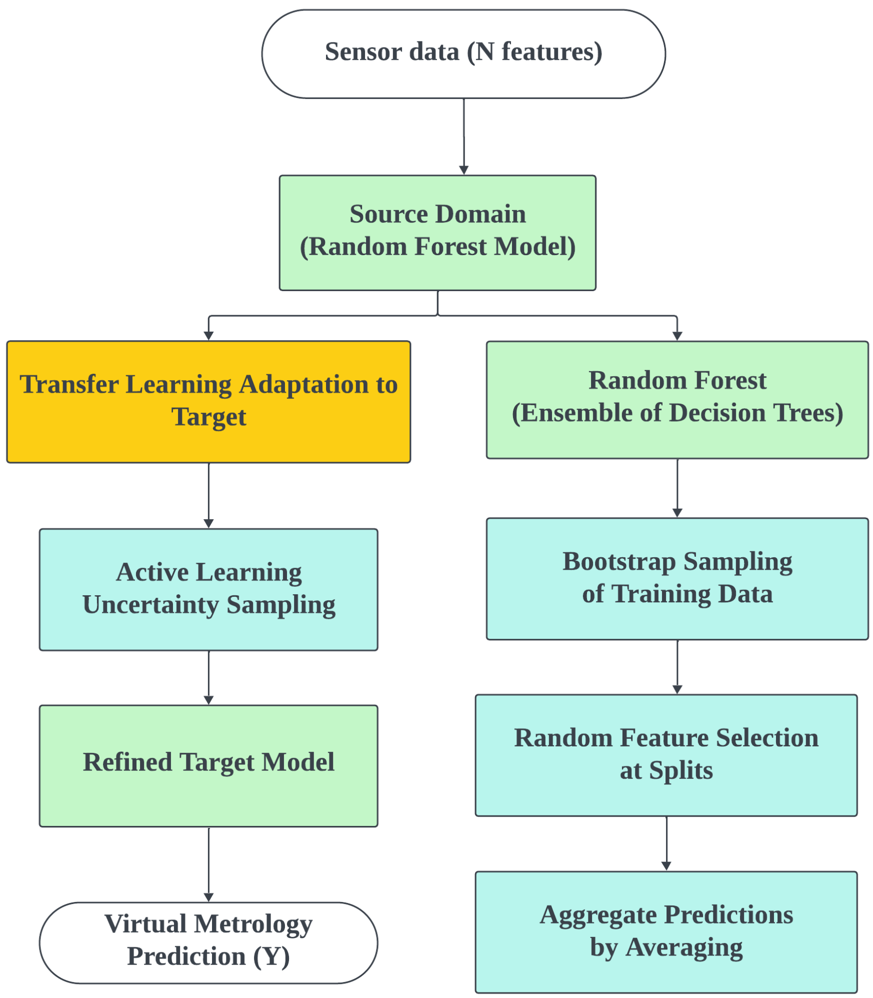

Figure1: Workflow for virtual metrology prediction using enhanced transfer learning with adaptation and active learning

The transfer learning architecture in this study is implemented in two main stages: adaptation and refinement. In the first stage, a base regression model is pre-trained using historical sensor data from the source domain, which may include multiple factories, process stages, and measurement types. This pre-trained model is then adapted to the target domain by fine-tuning with a limited set of labeled data specific to the target context. The adapted model leverages knowledge from the source domain to improve generalization and predictive accuracy under new manufacturing conditions. In the second stage, the adapted model undergoes iterative refinement through active learning. During each iteration, the model identifies the most uncertain predictions within the target domain and incorporates these samples into the training set. This targeted refinement enables the model to focus on challenging cases, further enhancing its accuracy and robustness. By combining adaptation and active learning-driven refinement, the proposed architecture ensures strong predictive performance even in data-scarce and highly variable manufacturing environments.

3.4 Experiment Setting

3.4.1. Design of Experiments

To evaluate the effectiveness of transfer learning and active learning in virtual metrology modeling, we designed experiments encompassing cross-factory, cross-stage, and cross-measurement adaptation scenarios. Sensor datasets were collected from multiple Seagate semiconductor wafer manufacturing facilities, representing diverse process stages and measurement types. For cross-factory adaptation, data from one factory was used as the source domain, while data from another factory served as the target domain, with particular attention given to cases where the target domain dataset was limited in size. This setup allows us to assess the capability of transfer learning to address data scarcity and domain shift. Similarly, cross-stage and cross-measurement experiments were conducted by designating specific process stages and measurement types as source and target domains, respectively. In each scenario, the adapted model was first fine-tuned with limited labeled data from the target domain, followed by iterative refinement using active learning to incorporate the most uncertain samples.

3.4.2. Evaluation Metrics

The performance of virtual metrology models was evaluated using three regression metrics: mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (R^2). These metrics were computed for models trained solely on the target domain, as well as for models adapted and refined through transfer learning and active learning. The improvement in each metric was calculated as the percentage reduction in error or increase in R^2, as shown below:

- $\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$

- $\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$
- $R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}i)^2}{\sum^{i} \quad = 1^{n}(y_i - \bar{y})^2}$

4. Result and Discussion

The efficacy of the VM-TALA framework was validated using a large-scale, real-world dataset sourced from a Seagate semiconductor manufacturing facility. This dataset includes multiple sensors spanning various process stages, measurement types, and factory sites. The figures below illustrate changes in MAE, RMSE, and R^2 for each target context, comparing the performance of the Adapted Model and the Refined Model. Three metrics are used to comprehensively evaluate model performance across all contexts: mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (R^2).

Context Encoded	Target MAE	Target RMSE	Target RÂ²	Adapted MAE	Adapted RMSE	Adapted RÂ²	Refined MAE	Refined RMSE	Refined RÂ²
stage_0_meas_C_fac_Y	2.633312	4.653801	-1.5578	0.238719	1.094417	0.865295	0.174198	0.495497	0.972273
stage_1_meas_C_fac_X	18.69353	47.66746	-0.06946	5.704344	24.41296	0.921084	3.940854	18.80294	0.948848
stage_2_meas_C_fac_X	40.52906	51.97169	-0.06982	13.98225	18.85622	0.8873	14.05569	19.28378	0.884428
stage_3_meas_D_fac_X	1338.373	1345.76	-92.3638	2.966299	4.877093	0.997551	3.450357	15.87809	0.971147
stage_3_meas_A_fac_X	58.07634	109.6704	-0.12906	10.70599	22.41142	0.735111	10.05285	25.57469	0.905536
stage_4_meas_B_fac_X	4.775721	11.72476	-0.06063	0.435776	0.755348	0.974252	0.787988	4.218209	0.933157
stage_4_meas_B_fac_Y	11.56094	160.8559	-0.14909	15.41632	119.3902	0.872103	13.70549	115.6785	0.665294

stage_5_meas_E_fac_X	1.901159	2.870282	-0.42468	0.34821	0.499007	0.957094	0.414709	1.202834	0.917568
stage_6_meas_C_fac_Y	13.89912	163.7359	-0.21043	3.111309	26.00234	0.736153	4.398212	60.87194	0.785499

Table 1: Model performance metrics for each target context

Our results shown in Table 1 indicate that the MAE for target models trained only on local data is consistently high, reflecting poor predictive performance due to data scarcity and domain shift. For example, in stage_3_meas_D_fac_X, the target MAE is over 1300, and in stage_1_meas_C_fac_X, it is nearly 19. After domain adaptation, the MAE drops dramatically in all contexts. For instance, in stage_0_meas_C_fac_Y, MAE decreases from 2.63 as target to 0.24 as adapted, and in stage_5_meas_E_fac_X, from 1.90 as target to 0.35 as adapted. This demonstrates that leveraging source domain knowledge significantly improves prediction accuracy, even when the target domain has limited data. Active refinement further reduces MAE in most cases, though the improvement is sometimes modest compared to the jump from target to adapted. For example, in stage_3_meas_A_fac_X, MAE drops from 58.08 as target to 10.71 as adapted and then to 10.05 as refined. In some contexts, such as stage_4_meas_B_fac_Y, the refined MAE at 13.71 is slightly lower than the adapted MAE at 15.42, but both are much higher than the target MAE at 11.56.

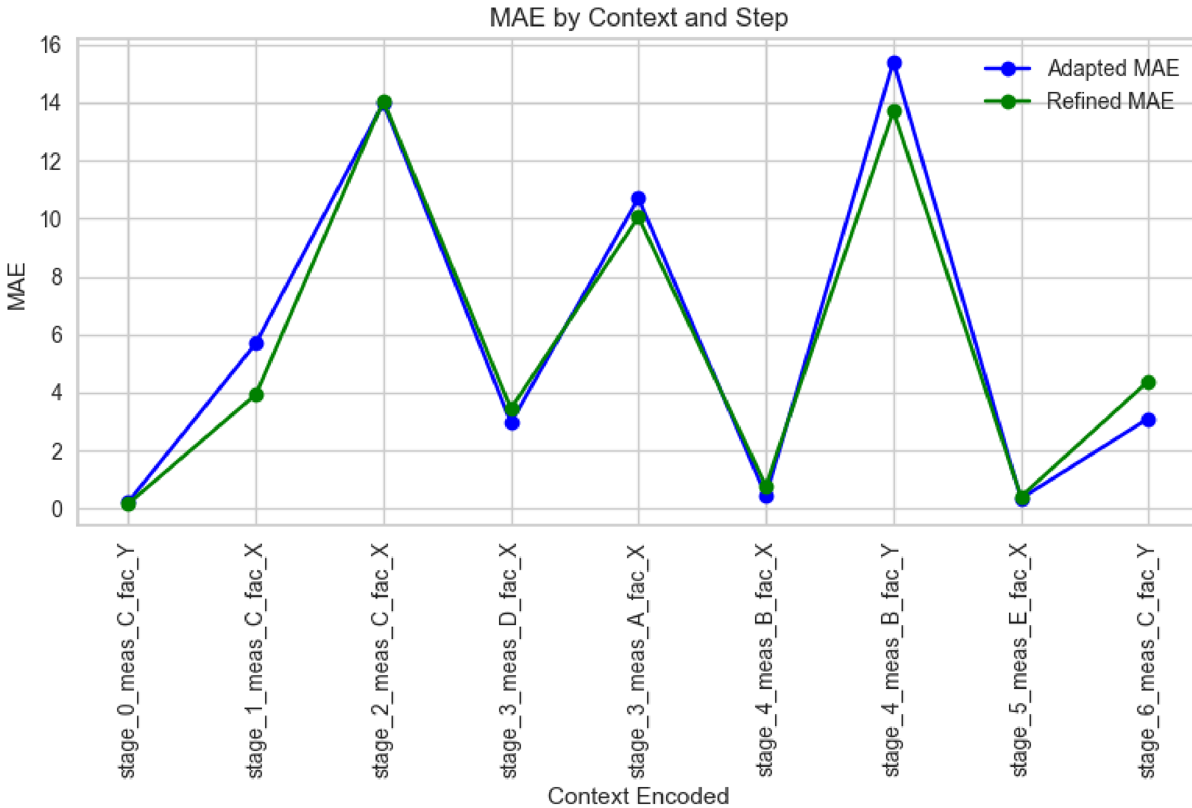

Figure 2: MAE drops dramatically in all contexts after active refinement

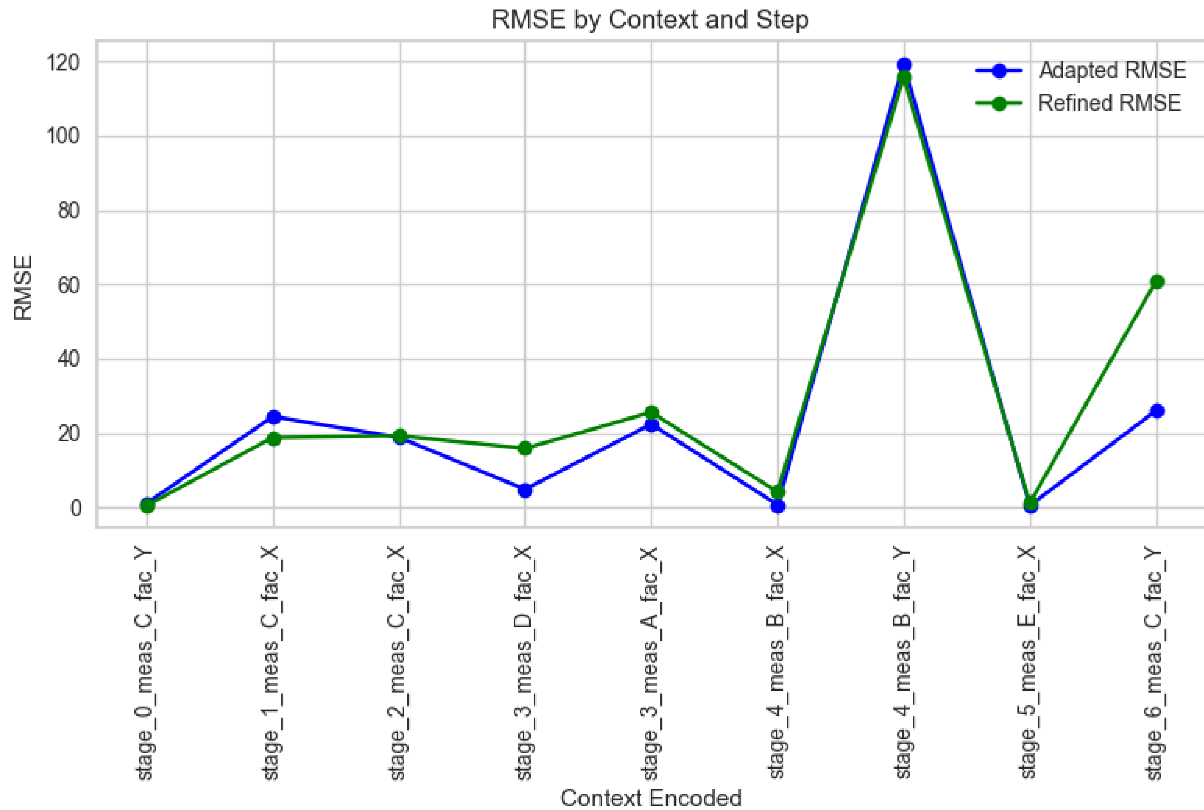

Figure 3 : Refined model shows indicate better model performance and more reliable predictions.

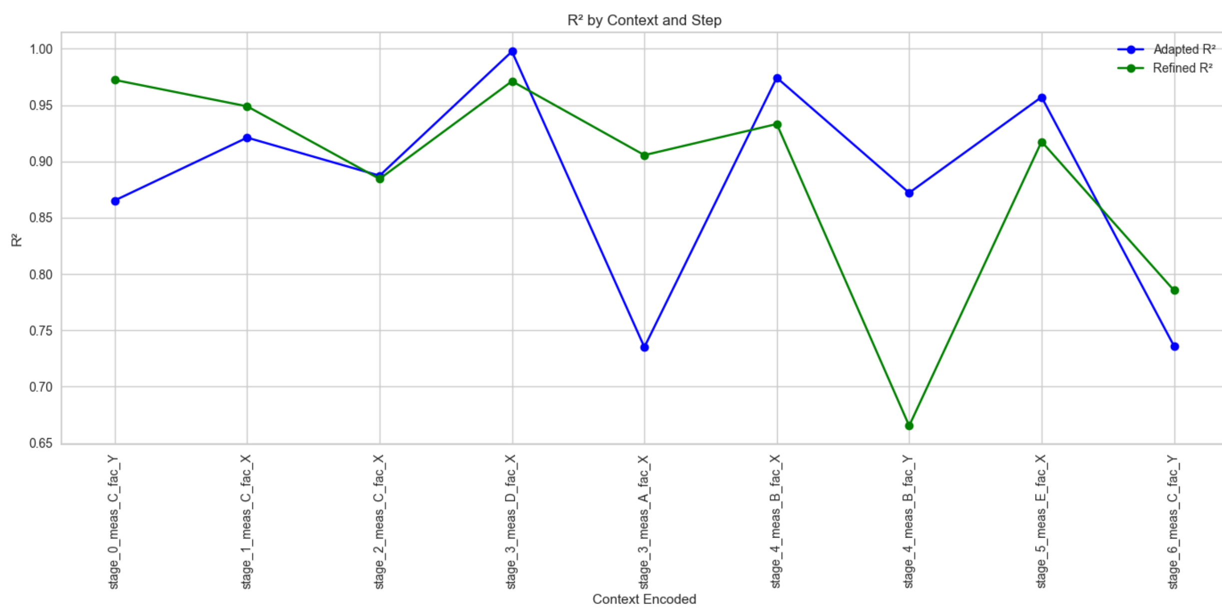

Figure 4: Active refinement further boosts R^2 for most of the context

Based on the insights drawn from Figure 3, it can be observed that the target models exhibit very high RMSE values, reflecting poor accuracy and instability in predictions. For instance, in stage_3_meas_D_fac_X, the target RMSE exceeds 1300, while in stage_1_meas_C_fac_X, it approaches 48. Such elevated RMSE values are characteristic of data-scarce domains with significant process drift, where models trained solely on local data struggle to generalize. Upon applying domain adaptation, RMSE is dramatically reduced across all contexts. For example, in stage_0_meas_C_fac_Y, RMSE decreases from 4.65 as target to 1.09 as adapted, and in stage_5_meas_E_fac_X, from 2.87 as target to 0.50 as adapted. This substantial reduction demonstrates that transfer learning effectively leverages source domain information to enhance prediction stability and accuracy. Active refinement further reduces RMSE in most cases, although the improvement is sometimes incremental compared to the initial adaptation. For example, in stage_3_meas_A_fac_X, RMSE decreases from 109.67 as target to 22.41 as adapted and then to 25.57 as refined. In other contexts, such as stage_0_meas_C_fac_Y, the refined RMSE reaches 0.50, which is lower than both the target and adapted RMSE, indicating enhanced model precision.

A similar trend is observed for the coefficient of determination, R^2. The target models frequently yield negative R^2 values, indicating poor fit due to data scarcity and domain shift. For example, in stage_3_meas_D_fac_X, the target R^2 is -92.36, highlighting the model's unreliability in that context. Following domain adaptation, R^2 improves markedly, often exceeding 0.7 or even 0.9, signifying that the adapted model can explain most of the variance in the target data. For instance, in stage_0_meas_C_fac_Y, R^2 increases from -1.56 as target to 0.87 as adapted. Active refinement further boosts R^2, with the refined model reaching values close to 1 in several contexts, such as 0.97 as refined in stage_0_meas_C_fac_Y. This progression underscores the refined model's high accuracy and robustness, even in challenging domains.

5.Conclusion

The results indicate that the proposed transfer learning and active learning framework is effective in improving both regression and classification performance in complex, multi-context manufacturing environments. Our experiments demonstrate consistently outperforms the baseline and adapted models, particularly in challenging contexts with initially poor performance. The refined model strategically incorporates additional data samples that exhibit high predictive uncertainty, as identified through active learning techniques. By focusing on these informative instances, the model can refine its understanding of complex patterns and improve generalization across diverse manufacturing contexts.

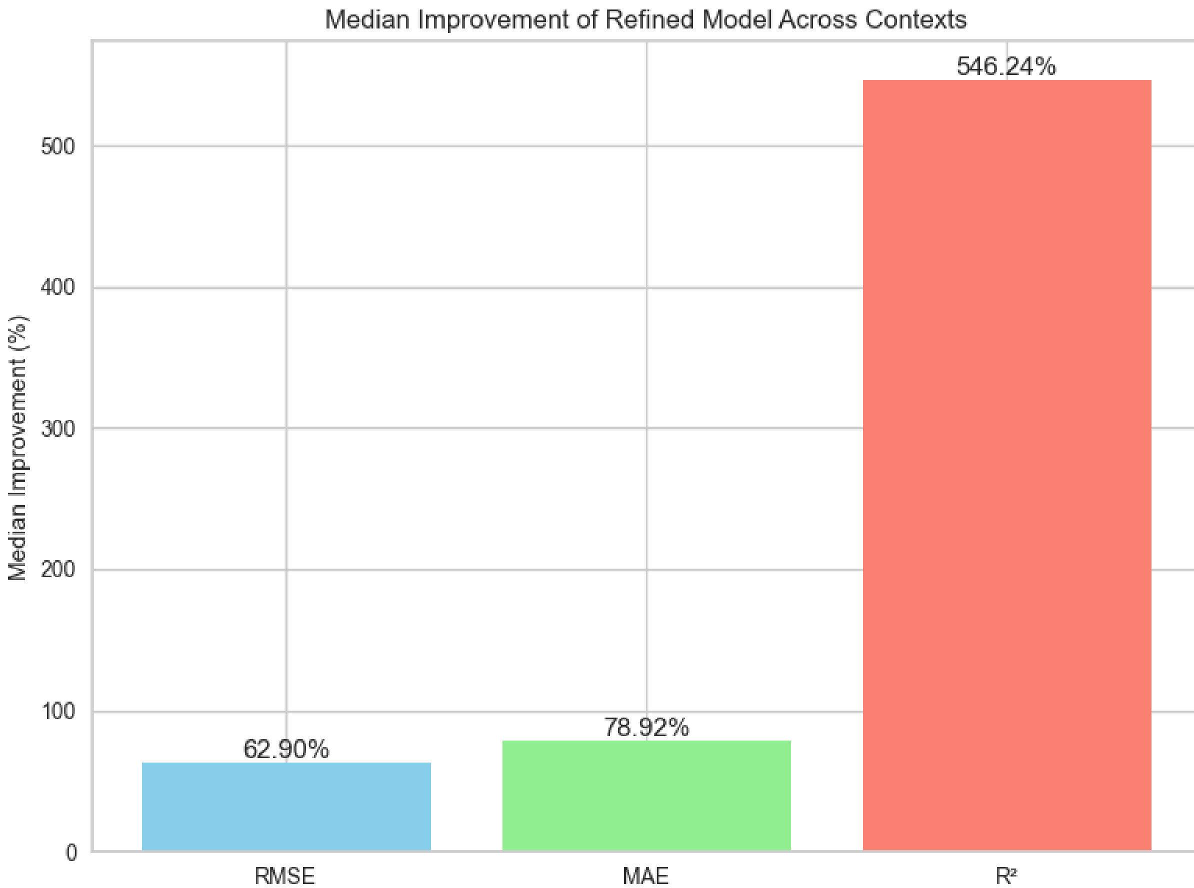

Figure 5: Median Improvement of Refined Model Across Contexts

The proposed VM-TALA architecture demonstrates substantial improvements in predictive accuracy and robustness across cross-factory, cross-stage, and cross-measurement scenarios. Notably, the refined model achieves a median improvement of 62.90% in RMSE and 78.92% in MAE. The refined model achieved its best performance with an R^2 of 0.97, a RMSE improvement of 1639.17%, and a MAE improvement of 99.74% across the evaluated contexts. By strategically selecting the most informative samples for model refinement, our approach addresses the challenges of data scarcity and limited diversity in target domains, enabling rapid and reliable deployment of VM models in high-mix, low-volume manufacturing environments. Overall, the VM-TALA framework streamlines the model adaptation process, reduces metrology overhead, and supports yield enhancement and cost reduction in dynamic industrial settings. Future work may explore the extension of this framework to multi-modal data sources and real-time adaptive learning, further enhancing its applicability in dynamic production environments. The findings underscore the potential of combining transfer learning and active learning to drive innovation in virtual metrology and predictive analytics for smart factories.

References

1. Bai, C., Dallasega, P., Orzes, G., & Sarkis, J. (2020). Industry 4.0 technologies assessment: A sustainability perspective. Journal of Cleaner Production, 276, Article 124087
2. Dreyfus, P.-A., Psarommatis, F., May, G., & Kiritsis, D. (2021). Virtual metrology as an approach for product quality estimation in Industry 4.0: A systematic review and integrative conceptual framework. International Journal of Production Research, 60(15), 4781-4806. https://doi.org/10.1080/00207543.2021.1976433
3. Guha, R., Suresh, A., DeFrain, J., & Deb, K. (2023). Virtual metrology in long batch processes using machine learning. International Journal of Production Research, 61(22), 7949-7965. https://doi.org/10.1080/00207543.2023.2220487
4. Kaitwanidvilai, S., Sittisombut, C., Huang, Y., & Bom, S. (2025). Optimizing product quality prediction in smart manufacturing through parameter transfer learning: A case study in hard disk drive manufacturing. *Processes, 13*(4), Article 962. https://doi.org/10.3390/pr13040962
5. Liang, Y., Wang, Y., Li, W., Pham, D. T., & Lu, J. (2025). Adaptive fault diagnosis of machining processes enabled by hybrid deep learning and incremental transfer learning. Journal of Manufacturing Systems.
6. Lin, C., Tseng, T., Hossain Emon, S., & Tsai, T. (2025). Large Pre-Trained Models and Few-Shot FineTuning for Virtual Metrology: A Framework for Uncertainty-Driven Adaptive Process Control in Semiconductor Manufacturing. *IEEE Transactions on Automation Science and Engineering*, *99*, 1-1.
7. Maitra, V., Su, Y., & Shi, J. (2024). Virtual metrology in semiconductor manufacturing: Current status and future prospects. Journal of Process Control, 131, Article 103290.
8. Semitela, Â., Pereira, M., Completo, A., Lau, N., & Santos, J. P. (2024). Improving Industrial Quality Control: A Transfer Learning Approach to Surface Defect Detection. Sensors, 24(2), Article 527. https://doi.org/10.3390/s24020527
9. Shoghi, R., Morand, L., Helm, D., & Hartmaier, A. (2024). Optimizing machine learning yield functions using query-by-committee for support vector classification with a dynamic stopping criterion. Computational Materials Science, 237, Article 112879.
10. Soori, M., Arezoo, B., & Dastres, R. (2023). Virtual manufacturing in Industry 4.0: A review. Journal of Manufacturing Systems, 68, 262-273.
11. Wang, L., Chen, S., Jiang, L., Pan, S., Cai, R., Yang, S., & Yang, F. (2025). Parameter-efficient fine-tuning in large language models: a survey of methodologies. Artificial Intelligence Review, 58(3), 227-258.
12. Ye, C., Yang, J., Kilic, Z. M., Chen, D., & Ding, H. (2024). Experimental modal transferring of industrial robot with data augmentation through domain adaptation and transfer boosting. Journal of Manufacturing Systems, 73, 202-212.
13. Zhu, Q., Lu, Z., & Hu, Y. (2024). Transfer learning-enhanced physics informed neural network for accurate melt pool prediction in laser melting. Journal of Manufacturing ,*2*(1), Article 0001.

Funding
This study was supported by Seagate Technology.

Data Availability Statement

The data presented in this study are available from the corresponding author upon request.

Conflicts of Interest

Authors Yu Huang and Sthitie Bom were employed by the company Seagate Technology. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.