2025 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)

Integrating Domain Adaptation and Causal Discovery in Digital Twins for Plastic Injection Molding

1st Gian Marco Paldino
Machine Learning Group
Université Libre de Bruxelles
Brussels, Belgium
gpaldino@ulb.ac.be

2nd Olivier Caelen
Sirris
Belgium

3rd Marouene Oueslati
Sirris
Belgium

4th Marc Ansay
Sirris
Belgium

5th Tojo Valisoa A. Johanesa
ILIA and GMECA Laboratories
University of Mons
Mons, Belgium

6th Gianluca Bontempi
Machine Learning Group
Université Libre de Bruxelles
Brussels, Belgium

***Abstract*—The development of digital twin (DT) technology is transforming the manufacturing industry, allowing for applications such as real-time process monitoring, predictive maintenance, and optimization of production systems. However, traditional DT frameworks often maintain a high-level perspective and do not consider the challenges arising from real industrial data. Factors such as changes in input materials and environmental conditions can hinder the functioning of DTs and are frequently overlooked. Using the example of the plastic injection molding industry, this paper highlights the necessity of including technologies such as domain adaptation and causal discovery in the development of DTs. Domain adaptation enables adaptation to changes, while causal discovery provides a deeper understanding of the underlying process dynamics. Their combined adoption allows DTs to achieve improved robustness and flexibility, extending their applicability across diverse manufacturing scenarios.**

***Index Terms*—Plastic Injection Molding, Digital Twin, Domain Adaptation, Causal Discovery, Machine Learning**

I. Introduction

Plastic injection molding is a cornerstone of modern manufacturing, vital for producing high-precision and intricate plastic components used across various industries. While the technology has seen significant advancements, ensuring consistent quality and operational efficiency remains a challenging endeavor due to the inherent complexities of the process. These challenges are compounded by the need for manufacturing systems to adapt to variable conditions and materials, necessitating innovations that can enhance adaptability and decision-making accuracy.

Gian Marco Paldino and Gianluca Bontempi are supported by the Service Public de Wallonie Recherche under grant nr 2010235–ARIAC by DigitalWallonia4.ai. Computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under Grant No. 2.5020.11 and by the Walloon Region.

Digital Twin (DT) technology, characterized by its ability to create a virtual replica of physical systems, presents a promising solution to these challenges [1] and has been used in various domains [2]. However, despite the potential, realization of DTs remains a challenging task [3], and most current implementations primarily propose high-level architectures without delving into practical applications using real-world industrial data [1], [3], [4]. This lack of practical implementation means that the actual challenges arising from real-world data—such as domain shifts due to variable conditions and materials—are often overlooked. Consequently, when these DT models are applied in practice, they encounter difficulties in maintaining accuracy and reliability under changing conditions, and domain adaptation techniques are often beneficial [5].

Moreover, in many industrial settings, a vast collection of variables is performed without a comprehensive understanding of the underlying processes. This abundance of data, without proper context or insight into causal relationships, hampers the ability to make informed decisions and effective interventions. Understanding which variables genuinely impact product quality and process efficiency is crucial for optimizing operations and achieving desired outcomes. Recent work [6] has shown that causal discovery can be effectively exploited to learn interpretable models of the smart factory environment through DTs. Moreover, [7] emphasized the role of causality in solving open problems in machine learning, such as transfer and generalization, by advocating for causal representation learning. This involves discovering high-level causal variables from observations and leveraging them to enhance transfer learning capabilities.

To address these gaps, this paper proposes the inclusion of domain adaptation and causal discovery in existing DT frameworks, to inherently tackle the challenges of domain shift and leverage causal discovery to understand the process

2766-8576/25/$31.00 ©2025 IEEE
DOI 10.1109/PerComWorkshops65533.2025.00050

at hand. By integrating domain adaptation, the DT can automatically adjust its models to account for variations in input materials and environmental factors, thus maintaining accuracy without the need for manual recalibration. Simultaneously, employing causal discovery allows for the identification and quantification of cause-and-effect relationships among process variables, enabling more informed and proactive adjustments and improving the quality of domain adaptation.

Within the DT framework, the learned models are continuously synchronized with sensor data from the physical injection molding setup. Domain adaptation permits the DT to remain accurate under shifts in materials or molds, while causal discovery confers interpretability when diagnosing variations. This combination strengthens the DT's ability to guide proactive process adjustments on the factory floor.

By using real-world industrial data, this paper aims to demonstrate its effectiveness in improving the precision, adaptability, and efficiency of the plastic injection molding process. This approach not only addresses the often-overlooked challenges inherent in practical implementations but also sets a new standard for smart manufacturing technologies.

The main contributions of this paper are as follows:

- We highlight the essential role of incorporating domain adaptation and causal discovery in the development of DTs and underline the importance of these techniques in enhancing operational efficiency.
- We validate the significant advantages of integrating domain adaptation and causal discovery by conducting rigorous experiments using real-world data from the plastic injection molding industry.

The remainder of the paper is structured as follows: Section II surveys the existing literature on DTs in manufacturing, particularly focusing on previous implementations in plastic injection molding and highlighting the innovation brought by integrating domain adaptation and causal discovery. Section III presents the data collection process and describes the dataset at our disposal. Section IV describes the methodology, detailing the specific techniques used for domain adaptation and causal discovery. Section V presents the results, discussing how our framework enhances predictive accuracy and process adaptability in real-world manufacturing settings. Section VI explores the broader implications and concludes the paper, summarizing the key findings and reiterating the importance of our contributions to the field of digital twins in industrial applications.

II. Related Work

Plastic injection molding is a complex manufacturing process involving numerous variables that affect the quality of the final product. Ensuring consistent quality requires a deep understanding of the process dynamics and the relationships among various parameters. In recent years, DT technology has emerged as a promising solution to model and simulate manufacturing processes, enabling real-time monitoring and optimization. However, existing approaches in DTs for injection molding often lack the integration of domain adaptation and causal discovery, limiting their adaptability and interpretability under varying operational conditions.

A. Digital Twins in Injection Molding

Several studies have focused on developing DTs for injection molding to enhance process control and product quality. [3] introduced a model-driven methodology for efficiently developing DTs of Cyber-Physical Production Systems (CPPSs) in injection molding. Their approach utilizes domain-specific languages (DSLs) to define reactive behavior and facilitate the setup of DTs, automating essential development activities such as Design of Experiment (DoE) execution and process optimization. Similarly, [8] aimed to transform a machine-variable controlled injection molding machine into a CPPS by augmenting it with a DT that predicts part quality from process variables. They employed a nonlinear state-space model to map process value trajectories to final part quality, highlighting the impact of environmental conditions and material properties on product quality. [9] proposed a methodological approach for optimizing the micro injection molding process by leveraging DTs for in-process monitoring. Their DT allows real-time monitoring of process parameters and comparison with analytical models to maintain optimal conditions. [10] developed SHION, an interactive DT supporting real-time shopfloor operations for smart thermoplastic injection. Their cloud-based DT incorporates AI-based control of process parameters to detect product failures in real time. [11] described an integrative industrial Internet architecture for the injection molding industry based on DTs. They established a smart factory architecture utilizing DTs for intelligent equipment, production lines, workshops, and factories. Finally, [12] discussed the data and technology required to build a DT model for smart injection molding. They emphasized the need for integrating fault detection systems and system integration to automate development activities. Their proposed DT model allows for fault detection, prognostic maintenance, and predictive manufacturing. While these studies contribute valuable insights into the development and application of DTs in injection molding, they generally focus on system integration, real-time monitoring, and automation. They do not explicitly address the challenges of domain adaptation—adjusting models to new materials, molds, or environmental conditions without extensive retraining. Additionally, they do not incorporate causal discovery to identify and leverage the causal relationships among process variables, which is crucial for interpretability and proactive quality control.

III. Data

The data collection campaign, realised at Sirris facilities, was designed using a Taguchi plan and a Design of Experiment (DoE) to capture a comprehensive array of control inputs and process variables essential for analyzing the plastic IM process. This dataset represents an extensive compilation of measurements, each corresponding to a unique production piece, thereby offering a granular view of the operational dynamics and quality outcomes of the molding process.

The dataset encompasses a wide variety of variables, categorized broadly into process parameters, environmental conditions, and quality outcomes. Process parameters include, but are not limited to, the velocities and pressures at different stages of the injection cycle, the temperatures across various zones of the heating cylinder, and the material flow rates.

This section provides a detailed overview of the data used in this study, which includes various drying experiments conducted under controlled conditions. The experiments involve different materials, drying times, and molds, aiming to capture the drying process's effect on injection molding quality. The data consists of time series records of process variables, which are analyzed to identify key interactions.

For coherence with the experimental results, we use the experiment names from the acquisition system using a unified naming space acquisition protocol. The productions are the following ones:

- **Production 65**: This production of 585 pieces uses ABS GP22 material with a rectangular test plate mold (Fig. 1a).
- **Production 87**: This production of 512 pieces uses ABS GP22 material with a rectangular test plate mold (Fig. 1a). It replicates the drying conditions of Production 65 but with more fine-grained drying intervals.
- **Production 88**: In this production of 554 pieces, ABS GP35 material is used with a rectangular test plate mold (Fig. 1a).
- **Production 89**: This production of 559 pieces uses PC/ABS Bayblend T65XF material with a rectangular test plate mold (Fig. 1a).
- **Production 97**: This production of 412 pieces uses ABS GP22 material with a tensile bar ISO A mold (Fig. 1b). It investigates the drying process using a different mold with the same ABS material as in Production 65 and 87.

(a) Rectangular test plate (b) Tensile bar ISO A

Fig. 1: Comparison of different molds used in the experiments

The dataset underwent thorough preprocessing to ensure its quality and consistency. Initially, features containing only missing values were removed, along with those that exhibited constant values across all rows, as they lacked meaningful information. Features not shared across all datasets were also excluded to maintain coherence. Any remaining rows containing missing values were subsequently removed to ensure each dataset was complete for analysis. Additionally, the initial batch of each production run was excluded to focus the analysis on relevant production data and avoid the influence of specific initialization dynamics.

After the aforementioned steps, the dimensions of the datasets are the following: Production 65, 497 pieces; Production 87, 437 pieces; Production 88, 498 pieces; Production 89, 510 pieces; Production 97, 340 pieces.

IV. Methodology

In this section, we present our methodology for integrating domain adaptation with causal feature selection to enhance the predictive accuracy and adaptability of the DT in plastic injection molding. We formalize the problem setting, introduce domain adaptation techniques, and detail how we use causal discovery in our domain adaptation methods to improve model performance while maintaining interpretability.

A. Problem Formalization

Let $c = (\text{mold type}, \text{material}) \in \mathcal{C}_{\text{tr}}$, where $\mathcal{C}_{\text{tr}}$ is the set of all combinations of mold types and materials in the training set. These configurations are part of a broader set $\mathcal{C}$ (i.e., $\mathcal{C}_{\text{tr}} \subset \mathcal{C}$).

Consider a batch $b_c^t = \{p_1^c, p_2^c, \ldots, p_{|b_c|}^c\}$ produced under configuration c with drying time t, where p_i^c is the i-th piece in the batch and $|b_c|$ is the batch size. A production run $\mathfrak{p}_c = \bigcup_t b_c^t$ encompasses all batches with varying drying times under configuration c.

For each piece p_i^c, we measure M continuous process variables $\mathbf{x}_i^c = \{x_{i1}^c, x_{i2}^c, \ldots, x_{iM}^c\}$. These variables follow a conditional distribution $f_{\mathbf{X}|C}(\mathbf{x} \mid c)$, indicating dependence on the configuration c.

Define the success label for batch b_c^t:

$$y_c = \begin{cases} 1, & \text{if } t > t_{\min}(c), \\ 0, & \text{otherwise}, \end{cases}$$

where $t_{\min}(c)$ is the minimum drying time required for success under configuration c. Each piece p_i^c in batch b_c^t inherits the label $y_i = y_c$.

The training dataset is:

$$\mathcal{D}_{\text{train}} = \{(c_i, \mathbf{x}_i^c, y_i)\}_{i=1}^N,$$

where N is the total number of pieces in the training set, c_i is the configuration, $\mathbf{x}_i^c$ are the measured variables, and y_i is the success label.

During inference on a new configuration $c' \in \mathcal{C} \setminus \mathcal{C}_{\text{tr}}$, we incorporate unsupervised data $\mathcal{D}_{\text{new}} = \{(c_i', \mathbf{x}_i^{c'})\}_{i=1}^{N'}$, where data points are sequentially observed, necessitating model adaptation.

B. Domain Adaptation Approaches

We propose three domain adaptation methods to address distributional shifts due to varying production conditions:

1) Adjustment with Standard Deviation: We adjust the source data to match the target domain statistics:

$$\widetilde{\mathbf{x}}_i = \left(\frac{\mathbf{x}_i^c - \widehat{\boldsymbol{\mu}}^c_{\text{source}}}{\widehat{\boldsymbol{\sigma}}^c_{\text{source}}} \right) \odot \widehat{\boldsymbol{\sigma}}_{\text{target}} + \widehat{\boldsymbol{\mu}}_{\text{target}},$$

where $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\sigma}}$ are the estimated means and standard deviations, and $\odot$ denotes element-wise multiplication.

2) Robust Adjustment: For robustness to outliers, we use median and interquartile range (IQR):

$$\widetilde{\mathbf{x}}_i = \left(\frac{\mathbf{x}_i^c - \widehat{\boldsymbol{\eta}}^c_{\text{source}}}{\widehat{\text{IQR}}^c_{\text{source}}} \right) \odot \widehat{\text{IQR}}_{\text{target}} + \widehat{\boldsymbol{\eta}}_{\text{target}},$$

where $\widehat{\boldsymbol{\eta}}$ is the median vector and $\widehat{\text{IQR}}$ is the IQR vector.

3) Adjustment with Covariance: Considering the covariance structure among variables, we align the data using Cholesky decomposition:

$$\widetilde{\mathbf{x}}_i = \widehat{\mathbf{L}}_{\text{target}} \widehat{\mathbf{L}}^{c^{-1}}_{\text{source}} (\mathbf{x}_i^c - \widehat{\boldsymbol{\mu}}^c_{\text{source}}) + \widehat{\boldsymbol{\mu}}_{\text{target}},$$

where $\widehat{\mathbf{L}}$ are lower triangular matrices from the Cholesky decomposition of the covariance matrices $\widehat{\Sigma}$.

C. Causal Feature Selection Using NOTEARS

To incorporate causality efficiently, we perform causal feature selection using the NOTEARS algorithm [13] and the concept of Markov Blanket (MB) [14]. Specifically, we apply the NOTEARS algorithm to $\mathcal{D}_{\text{train}}$ to learn a causal graph G, represented as a Directed Acyclic Graph (DAG). Nodes correspond to variables $\{X_1, X_2, \ldots, X_M, Y\}$, and edges represent causal relationships.

The NOTEARS algorithm formulates structure learning as a continuous optimization problem:

$$\begin{aligned} \min_{W} \quad & \frac{1}{2N} \sum_{i=1}^{N} \left\| \mathbf{x}_i - W^\top \mathbf{x}_i \right\|_2^2 + \lambda \|W\|_1 \\ \text{s.t.} \quad & h(W) = 0, \end{aligned} \tag{1}$$

where $W \in \mathbb{R}^{M \times M}$ is the weighted adjacency matrix, λ is a regularization parameter, $h(W) = \text{tr}\left(e^{W \circ W}\right) - M = 0$ ensures acyclicity, $\circ$ denotes the Hadamard (element-wise) product and $e^{W \circ W}$ is the matrix exponential.

From the learned causal graph G, we identify the Markov blanket $\mathcal{M}_Y$ of the target variable Y, which includes the parents of Y, denoted $\text{Pa}(Y)$, the children of Y, denoted $\text{Ch}(Y)$, the parents of Y's children (excluding Y), denoted $\text{Pa}(\text{Ch}(Y)) \setminus \{Y\}$. By selecting features in $\mathcal{M}_Y$, we focus on variables that have direct causal influence on Y or are directly influenced by Y, $\mathbf{x}_i^{\mathcal{M}_Y}$, and we apply the domain adaptation methods described earlier specifically to $\mathbf{x}_i^{\mathcal{M}_Y}$. This focuses the adaptation on causally relevant features, enhancing model robustness and interpretability.

While NOTEARS is a powerful structure-learning approach, we acknowledge that purely observational data and unobserved confounders can limit guarantees of true causality [15]. Nonetheless, despite the proposed framework is designed to seamlessly accommodate any causal discovery algorithm, adopting NOTEARS provides interpretable structures that guide feature selection and domain adaptation.

D. Experimental settings

To simulate the scenario of encountering a new production configuration not present in the training data, we adopt a leave-one-production-out cross-validation approach. This approach ensures that the test data originates from a production configuration (combination of mold type and material) not seen during training, thereby evaluating the model's ability to adapt to new domains. For all experiments, we use a Logistic Regression classifier as the predictive model,implemented using the scikit-learn library with default parameters. The NOTEARS algorithm is implemented using an open-source implementation, with the default regularization parameter λ. To evaluate the performance of the models, we compute the F1 Score, the harmonic mean of precision and recall, providing a balance between false positives and false negatives following metrics on the test set. Code to replicate the adaptation and causal discovery pipeline will be shared upon request.

V. Results and Discussion

We begin by presenting the results of the causal discovery step. Anticipating that the same causal structures would manifest across different production batches, we applied the NOTEARS algorithm separately to each batch. This approach was chosen to avoid mixing distributions, which could violate the independent and identically distributed (i.i.d.) assumption fundamental to causal discovery algorithms. Weak connections were filtered out by setting the weight threshold to 1. To further refine our selection, we included only those edges that appeared in more than half of the productions, utilizing a majority voting scheme. This process resulted in the graph presented in Figure 2, which illustrates the Markov blanket of Drying Time.

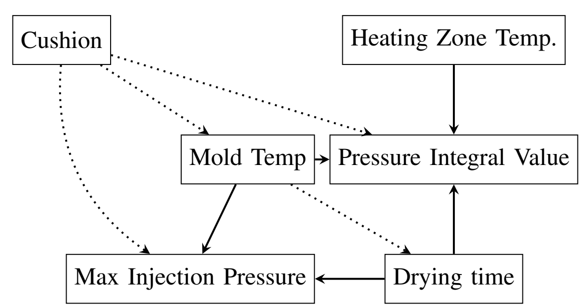

Fig. 2: Markov blanket of Drying Time. Detected causal connections that do not adhere with domain knowledge are reported with dotted lines.

Variations in drying time affect the moisture level, which in turn impacts the material viscosity and flow behavior, thereby influencing both the integral value of the pressure curve and the maximum injection pressure. The temperature of the heating area also affect the integral value of the

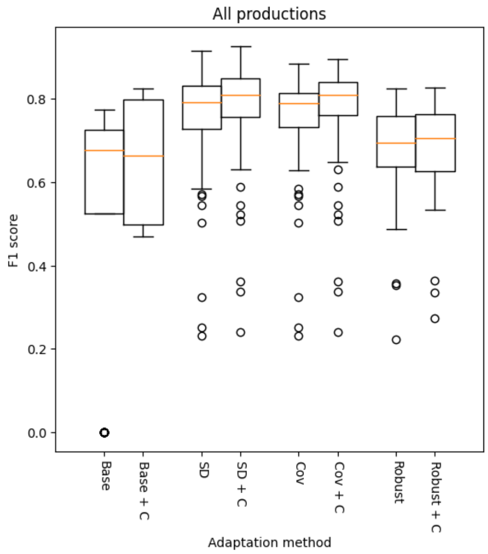

Fig. 3: Aggregated F1 scores distributions for various adaptation methods.

pressure curve. While most connections in this graph appear appropriate, we encounter a challenge with the causal links originating from the material cushion. According to domain knowledge, the cushion is a consequence of process conditions and does not causally influence other parameters. To reconcile this discrepancy, we recognize that the statistical associations identified by the algorithm may be due to common underlying factors influencing both the cushion and these variables, rather than direct causation from the cushion. Another unexpected causal connection is the one from Mold Temperature to Drying Time. However, since Mold Temperature affects the Pressure Integral value, it remains part of the Markov Blanket. We expect these relationships to be reversed. Nevertheless, this would not affect the Markov Blanket and the consequent considerations.

Figure 3 presents the aggregated results of the F1 score distributions across various domain adaptation methods, depicted using boxplots for all production scenarios. The Base approach corresponds to the absence of any adaptation method. The SD method, denoting adaptation using standard deviation, demonstrates the best performance among the methods tested, closely followed by the Cov method, which represents adaptation via covariance. The Robust method, utilizing the median and interquartile range (IQR) for adaptation, exhibits significantly poorer performance compared to the others. Methods labeled with '+C' indicate that causal feature selection was applied prior to performing the adaptation. Each method adjusts to changes in environmental conditions and material variations. In the following analysis, we focus on the SD approach to compare the effects of causal feature selection.

Figure 4 delineates the outcomes of our domain adaptation experiments across different production scenarios: 65 (above), 87 (center), and 88 (below). Given the consistency of the findings with Productions 89 and 97, results for the latter two are omitted for brevity. Each experiment was replicated 20 times to calculate the means and standard deviations, which are depicted in the figure.

The x-axis quantifies the number of data points from the target domain available for adaptation, while the y-axis measures the F1 Score.

For Production 65, it is observed that the absence of adaptation significantly hampers performance, whereas the implementation of any adaptation strategy markedly enhances it. Intriguingly, employing causal feature selection alone yields the highest performance for Production 65, although this strategy does not exhibit similar efficacy across other productions. For Production 87, the introduction of causal feature selection alone does improve performance over the baseline; however, the combination of adaptation strategies, both with and without causal feature selection, surpasses all baselines. Conversely, for Production 88, causal feature selection alone detrimentally impacts performance. Nonetheless, integrating standard deviation adaptation with causal feature selection not only mitigates this drawback but also achieves the highest performance.

These models, incorporating domain adaptation, consistently outperform those without such adjustments, which is critical when addressing distributional shifts due to variations in material types, mold geometries, or environmental conditions. This improvement manifests in enhanced accuracy and robustness of model predictions, emphasizing the necessity to tailor models to align with the statistical properties of new domains. Moreover, causal features provide valuable insights into the underlying process mechanisms, enabling the model to render more informed predictions even under novel or previously unseen configurations.

VI. Conclusion and Future Work

The integration of domain adaptation and causal discovery into digital twins for plastic injection molding marks a significant advancement in the field of smart manufacturing. This approach enhances the adaptability and predictive accuracy of digital twins, enabling them to maintain high performance under varying materials, molds, and environmental conditions. It addresses a critical challenge in industrial applications, where models often struggle with distributional shifts, leading to reduced reliability and efficiency.

While we demonstrate these methods in plastic injection molding, the same domain adaptation and causal discovery framework can be employed in other industrial processes (e.g., metal casting, semiconductor manufacturing) that face similar distribution shifts and require interpretable, data-driven models.

A critical direction for future work involves transitioning from static causal discovery methods to dynamic approaches that inherently account for the temporal structure in data. Time series–specific algorithms can capture how causal relationships evolve over time, providing richer insights into lagged interactions and phase shifts. Notably, the approach detailed in [16], [17]—which computes asymmetric conditional mutual information within Markov blankets and frames causal discovery as a supervised learning problem—offers a promising roadmap for handling such time-dependent dynamics. In conclusion, this work demonstrates that combining domain adaptation with causal discovery significantly enhances the robustness and flexibility of digital twins, offering a promising pathway toward more intelligent and adaptive manufacturing systems.

REFERENCES

[1] Y. Lu, C. Liu, I. Kevin, K. Wang, H. Huang, and X. Xu, "Digital twin-driven smart manufacturing: Connotation, reference model, applications and research issues," *Robotics and computer-integrated manufacturing*, vol. 61, p. 101837, 2020.

[2] G. M. Paldino, F. De Caro, J. De Stefani, A. Vaccaro, D. Villacci, and G. Bontempi, "A digital twin approach for improving estimation accuracy in dynamic thermal rating of transmission lines," *Energies*, vol. 15, no. 6, p. 2254, 2022.

[3] P. Bibow, M. Dalibor, C. Hopmann, B. Mainz, B. Rumpe, D. Schmalzing, M. Schmitz, and A. Wortmann, "Model-driven development of a digital twin for injection molding," in *International Conference on Advanced Information Systems Engineering*. Springer, 2020, pp. 85–100.

[4] F. Tao and M. Zhang, "Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing," *IEEE access*, vol. 5, pp. 20 418–20 427, 2017.

[5] G. M. Paldino, F. De Caro, J. De Stefani, A. Vaccaro, and G. Bontempi, "Transfer learning-based methodologies for dynamic thermal rating of transmission lines," *Electric Power Systems Research*, vol. 229, p. 110206, 2024.

[6] M. Lippi, M. Martinelli, M. Picone, and F. Zambonelli, "Enabling causality learning in smart factories with hierarchical digital twins," *Computers in Industry*, vol. 148, p. 103892, 2023.

[7] B. Schölkopf, F. Locatello, S. Bauer, N. R. Ke, N. Kalchbrenner, A. Goyal, and Y. Bengio, "Toward causal representation learning," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 612–634, 2021.

[8] A. Rehmer, M. Klute, H.-P. Heim, and A. Kroll, "A digital twin for part quality prediction and control in plastic injection molding," in *Modeling, Identification, and Control for Cyber-Physical Systems Towards Industry 4.0*. Elsevier, 2024, pp. 79–109.

[9] G. E. Modoni, B. Stampone, and G. Trotta, "Application of the digital twin for in process monitoring of the micro injection moulding process quality," *Computers in Industry*, vol. 135, p. 103568, 2022.

[10] F. J. Lacueva-Pérez, S. Hermawati, P. Amoraga, R. Salillas-Martínez, R. del Hoyo-Alonso, and G. Lawson, "Shion (smart thermoplastic injection): An interactive digital twin supporting real-time shopfloor operations," *IEEE Internet Computing*, vol. 26, no. 3, pp. 23–32, 2020.

[11] Z. Wang, W. Feng, J. Ye, J. Yang, and C. Liu, "A study on intelligent manufacturing industrial internet for injection molding industry based on digital twin," *Complexity*, vol. 2021, no. 1, p. 8838914, 2021.

[12] S. Nasiri, M. R. Khosravani, T. Reinicke, and J. Ovtcharova, "Digital twin modeling for smart injection molding," *Journal of Manufacturing and Materials Processing*, vol. 8, no. 3, p. 102, 2024.

[13] X. Zheng, B. Aragam, P. K. Ravikumar, and E. P. Xing, "Dags with no tears: Continuous optimization for structure learning," *Advances in neural information processing systems*, vol. 31, 2018.

[14] D. Koller, "Probabilistic graphical models: Principles and techniques," 2009.

[15] M. Kaiser and M. Sipos, "Unsuitability of notears for causal graph discovery when dealing with dimensional quantities," *Neural Processing Letters*, vol. 54, no. 3, pp. 1587–1595, 2022.

[16] G. Bontempi and M. Flauder, "From dependency to causality: a machine learning approach." *J. Mach. Learn. Res.*, vol. 16, no. 1, pp. 2437–2457, 2015.

[17] G. Bontempi, "Learning causal dependencies in large-variate time series," in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–7.

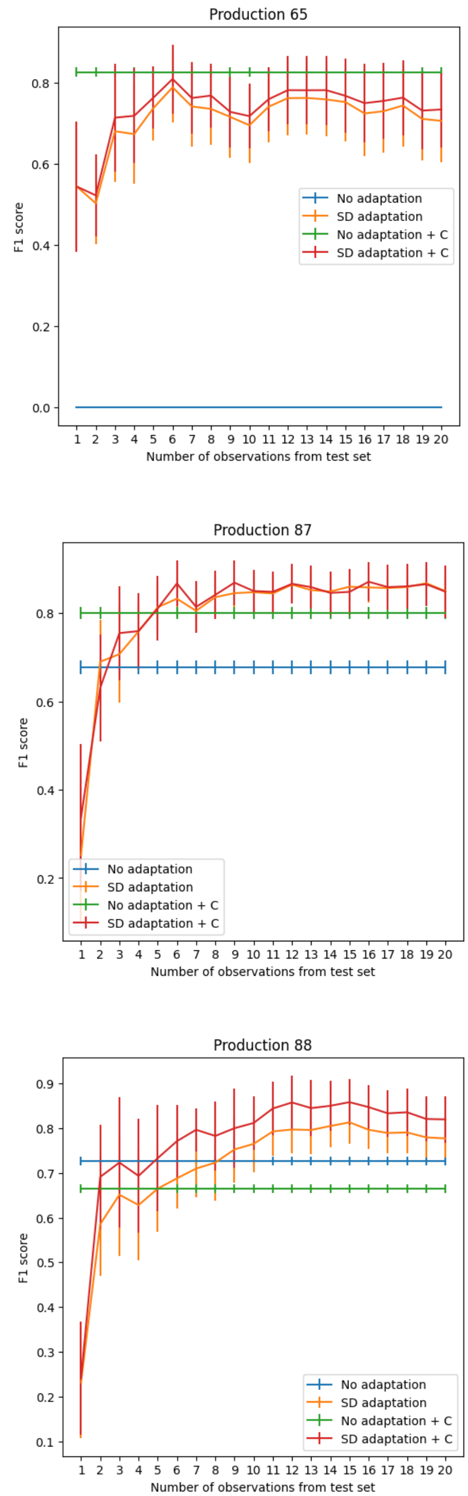

Fig. 4: F1 score per domain adaptation technique over 20 repetitions of random test point selection for Production 65, 87, 88. The x-axis represents the number of available points from target domain to perform the adaptation.