Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 200 (2022) 1550–1564

Procedia Computer Science

www.elsevier.com/locate/procedia

3rd International Conference on Industry 4.0 and Smart Manufacturing

Identifying key interactions between process variables of different material categories using mutual information-based network inference method

Shailesh Tripathi [a,*], Herbert Jodlbauer[a], Christian Mittermayr[b], Frank Emmert-Streib[c,d]

[a] *Production and Operations Management, University of Applied Sciences Upper Austria, Austria*
[b] *Greiner Packaging International GmbH, Greinerstrasse 70 4550 Kremsmuenster*
[c] *Predictive Society and Data Analytics Lab, Tampere University, Finland*
[d] *Institute of Biosciences and Medical Technology, Tampere, Finland*

Abstract

This paper analyzes production data from injection molding processes to identify key interactions between the process variables from different material categories using the network inference method called "bagging conservative causal core network" (BC3net). This approach is an ensemble method with mutual information that is measured between process variables to select pairs that show significant shared information. We construct networks for different time intervals and aggregate them by calculating the proportion of significant pairs of process variables (weighted edges) for each production process over time. The weighted edges of the aggregated network for each product are used in a machine learning model to optimize the network interval size (interval split) and feature selection, where edge weights are the input features and material categories are the output classification labels. The time intervals are optimized based on the classification accuracy of the machine learning model. Our analysis shows that the aggregated edge features of inferred networks can classify different material categories and identify critical features that represent interdependence in the associated process variables. We further used the "one vs. other" labels for the machine learning models to identify material-specific interactions for each material category. Additionally, we constructed an aggregated network over all samples in which the process variable interactions were steady over time. The resulting network showed modular characteristics where process variables of similar categories were grouped in the same community.

Keywords: injection molding, process variable interactions, process variable network, network inference, machine learning models.

* Corresponding author.
E-mail address: shailesh.tripathi@fh-steyr.at

10.1016/j.procs.2022.01.356

1. Introduction

Injection molding is a common cost-effective method to produce large-scale, simple-to-complex, identical plastic products. In injection molding production processes, a melted polymer is injected into a mold cavity, packed under pressure, and cooled until it has sufficiently solidified into the required shape. During production, the material flow, mold design, and dynamics of the machine parts are controlled by appropriate processing parameters to ensure quality production. During the injection molding production process, various process parameters for the injection molding machinery are tuned with each other to determine the optimal combination to ensure the plastic product quality [5]. In a complex manufacturing system, process variables and their interactions have a significant impact in determining product quality. Therefore, controlling/tuning individual process variables is insufficient for optimizing product quality and should consider the interactions between process variables to control production processes effectively and ensuring product quality.

There are various processing parameters in injection molding that require careful attention for consistent machine control to maintain production quality. A lack of consistent machine control causes material degradation and inferior product quality that is discarded as scrap [19]. There are different types of quality problems such as shrinkage, warpage, color and burn marks, surface texture quality, shape distortion, and other aesthetic defects [19, 16]. In real-world industrial production, maintaining consistent product quality requires a detailed domain understanding of parameters dependencies; otherwise, the product quality can be inconsistent between production lots for various reasons. These variations can be due to different process settings, including machine-, product-, and material-related parameters and should be controlled and monitored throughout the production process.

The production outcome of plastic products in injection molding can be described as complex due to the interdependence between several process variables that are neither random nor regular and depend on various material characteristics and external effects. Thus, the production output quality can be determined from various sets of causal interactions between process parameters in different cases [10, 4, 25]. Neglecting interactions between process variables leads to inefficient optimization of their settings. Interdependence of process variables can be put into two categories. First is the general interdependence as observed in all production types. Second is material-specific interactions that require material-specific settings. The interdependence of process variables can vary over time as the production process goes through different steps, each of which requires various parameter settings over time. The relationships between process variables can be estimated using various statistical and machine learning models that utilize process variable data from different production scenarios. Analyzing the interdependence between process variables through static correlation and association measures between pairs of process variables ignores the time-dependent effects of the interactions. Here, we consider the time-dependent effects of different production scenarios with various material categories (shapes and material type), parameter settings, machine types, and several other external factors that are sensitive to the process control and production output. The underlying concept of process variable data analysis is to identify critical parameters with material-specific dependencies between various process parameters over time and utilize that information to optimize the process control, product quality, and maintenance-related issues.

We analyze data from real-world injection molding processes consisting of 58 process variables and 215 product types over 8 material categories. Our analysis utilizes the ensemble-based network inference method based on bagging to find the linear and non-linear dependencies using the mutual information between process variables of products from different categories over various time intervals. In addition, we perform time-wise aggregation of inferred product networks. The network aggregation for each product is represented as edge weights of the network at different time intervals (proportion of edges show significant associations between two process variables over time). The edge weights for the inferred networks of each product are used as features to classify material categories and feature selection for the 8 material categories. We then apply the generalized linear model with regularization (elastic net penalty) [37, 11] to classify material categories with 10-fold cross-validation to select the best subset of features (interactions) for each material category. We further perform overall network aggregation across all products to estimate steady interactions over time as a network edge for a product that is consistent over all intervals. Finally, the overall network aggregation is utilized to group process variables into different categories.

This approach is divided into three steps. The first step is to infer networks for different time intervals using an ensemble-based network inference method called “bagging conservative causal core network” (BC3net) and to optimize the interval-split (dividing data into different intervals) using the Gernalized Linear Model (GLM) for multi-class classification over the 8 material categories and select an interval split as it has the lowest prediction error. The second step is to use the best interval split (minimum prediction error) for feature selection (interactions) over each material category using the GLM with regularization for the "one vs. others" two-class classification model. The third step is the overall network aggregation to group process variables that show steady interactions over time. The primary objective of this paper is to identify the material-specific interdependence of process variables using time-interval-based network inference. The secondary objective is to identify groups of process variables that are steady over time and utilize them for further product quality optimization. The tertiary objective is to find the material-specific interactions between process variables from Refs. [31, 30], which are identified as key features to estimate the scrap rate for similar production cases. Thus, this paper contributes to the application of network-based and machine learning methods to identify key material-specific interactions between process variables that play important roles in the production quality.

The structure of this paper is as follows. Section 2 provides a brief overview of related works. Section 3 describes the details of the data preprocessing and network inference methods, overall aggregation, and feature selection. Section 4 presents the results of the analyses and compares the interactions of process variables used as input variables in other studies. Finally, Section 5 provides concluding remarks about the results of our analysis.

2. Literature Review

Various studies have been utilized to address quality optimization in injection molding processes. Many studies related to quality optimization are based on Taguchi experimentation with fewer key process variables that are responsible for product quality [28, 27, 32, 5, 3, 22, 6, 20, 15]. Several computational approaches have been considered to optimize product quality, which use gradient-based approaches, evolutionary algorithms, and mixed approaches [35, 36, 20, 34, 6]. Reviews of frameworks that optimize injection molding methods are described in [16, 8, 26, 12].

[24] applied Support Vector Machine (SVM) [7] models by tuning different hyperparameters for error classification. The required data for the analysis were generated with an experimental set up on a *Demag* injection molding machine with *Hostacom DM2 T06* polymer and the *DN502* mold to predict product quality. The classification models *cycle time*, *dosage time*, *injection time*, *cushion*, *peak melt temperature*, and *ram velocity* were used as input variables. The response variables described as product defects were classified into 6 classes as Streaks, Strains, Burn marks, Edges, Unfilled parts, and Warped parts. [18] performed feature learning and process monitoring by utilizing deep learning with a convolution-deconvolution autoencoder to predict product quality. The experimentally-generated data were used on a *JSW J110ADC-180H* electric injection molding machine by tuning different process conditions. The input process variables, *screw displacement*, *injection pressure*, and *cavity pressure* were supplied in a 4D input tensor as $\chi(B \times V \times T \times C)$, where B is the batch size of a product quality class, V is the number of variables, T is the set of time instances, and C is the number of feature channels. [31] utilized various statistical features for 65 process variables using beta and SVM regression models to predict the scrap rate in production processes. The data were collected from the real-time production of products with different shapes, sizes, and material types as produced by 33 different machines within the company. [13] performed a detailed analysis of products and process-related variables as correlated with product quality to develop product and process fingerprints for quality monitoring systems. Various network-based approaches were used to address manufacturing-related problems that are relevant for injection molding machine process variable data analyses [17]. [33] used partial mutual information to infer the directed network to identify the causal effects between quality characteristics of the manufacturing process. [23] used normalized mutual information correlation analyses and network deconvolution for feature selection (parameters) as correlated with the power of the engine production system. The ensemble method of the BC3net [9, 1], was applied to infer networks for gene expression biological data. A similar approach was applied in financial data analysis [2] to find the causal interactions within investor networks.

3. Methodology

This section provides a brief overview of the methods applied for preprocessing, network inference, and feature selection. We first describe the available data for our analysis and the major preprocessing steps. We then extract relevant process variables for the subsequent network inference method after the data are split into different production lots (segments). The objective is to infer networks for different time intervals over each product to obtain information about process variable interactions that influence the production of a material category. In addition, the optimization of the material category-specific interaction features is useful to determine the interdependence of the process variables, which are used primarily to predict product quality.

3.1. Data collection

We collect raw data from 31 injection molding machines. These data sets are recorded in different files over a given production process. Additionally, we use the data from the enterprise resource planning (ERP) system. The relevant data for each production process are recorded in a production file, process data files, and an export file from the ERP system, which have the following information:

Production file: provides timestamps, cycle counter, tool-name, raw-material information, cavities, and set cycle time.
Process data: provides timestamps, cycle counter, set cycle time, and 90 different process variables such as the actual cycle time, temperature, pressure, volume, position, and rotational speed.
ERP data: provides the order number, material number, number of produced parts, and count of scrap parts.
Product information data: provides the machine number, material number, and material categories.

The production and process data files contain information from a period when multiple product types are produced. The process data do not contain product-type information. We use the production and ERP data to split the process data into different segments based on the product type and order number. The segments based on the product type and order number are described as process segments. Each process segment contains information from approximately 90 different process variables for the production process of a product type between the start and end times of the production order. The segment data file is a multivariate file where the columns are the process variables and the rows are their respective measures at different times. Relevant metadata about each production order are collected in a separate master data file, which includes the start and end times of production, order number, machine number, raw material number, and total production output in the units. We additionally receive material category information for different material types.

3.2. Data selection and filtering

We assume that there are 31 production machines described as $m = m_1, m_2, \ldots m_n$, with X being a set of r process variables as $X(m_i) = \{X_1, X_2, \ldots X_r\}$. Each $X(m_i, p_j)$ records values during the production process of product p_j at machine m_i between the time points t_1 and t_n, which is defined as $X(m_i, p_j, r_k) = (x_{t_1}, x_{t_2}, \ldots, x_{t_n})^T$. We extract $|r| = 58$ process variables for each machine m_i of the controller version CCXXX during the production of product type p_j. We found that the readings of 58 variables are not present in all the data, so we select process variables that are present in 80% of the data. Thus, we select 40 process variables and data from production processes that run for at least 6 to 120 h. Our analysis filters out process segments that run longer than 120 h and shorter than 20 min. This information is extracted using the master data file. Additionally, we filter out the process variables that do not show variability and have a large proportion of missing values for different machines and material types (product). Thus, for 20 machines and 97 material types, we have 215 unique data sets with 40 common process variables from 8 material categories. The product labels and their count is shown in Table 1.

Table 1. Data from 215 production processes over 8 material categories.

Product labels	ABS	PA	PC	PE	PET	POM	PP	PS
Frequency	32	59	17	12	7	16	63	9

3.3. Network inference method

The BC3net-like approach to infer a network is described in the following six steps. Step 1: The first step of the analysis of the BC3net approach is to test the following null hypothesis by performing mutual information tests between each pair of process variables using the bootstrap data set of X:

$$H_0 : I(X_i, X_j) = 0$$

$$H_1 : I(X_i, X_j) \neq 0$$

The mutual information is measured as follows [21]:

$$I(X_i, X_j) = -\frac{1}{2} log(1 - \rho^2).\ \rho \text{ is the correlation measure.}$$

Step 2: In each data set, the $\frac{n(n-1)}{2}$ hypotheses are simultaneously tested and obtained $p(i, j)$ value for each pair. To control the false discovery rate (FDR), multiple testing correction (MTC) [14] is applied, we adjust $p(i, j)$ to $p_{adjusted}(i, j)$, and defines a graph based on the adjacency matrix using the significance of the mutual information between each pair after applying the MTC as follows:

$$G_b(i, j) = \begin{cases} 1 \text{ if } p_{adjusted}(i, j) < \alpha \\ 0 \text{ otherwise} \end{cases}$$

Step 3: Once the $\{G_b\}_{b=1}^{B}$ is constructed, we filter the edges of each G_b by selecting one edge among all others for each process variable that has maximum mutual information. Thus, a $\hat{G}_b \subset G_b$ is selected for each bootstrapped data set and is repeated $B = 1000$ times.
Step 4: The next step is the aggregation of the Bootstrap data set where the aggregated weighted network $\{\hat{G}_b\}_{b=1}^{B=1000} \longrightarrow G_w$ is constructed. The weights are the proportional pairs of process variables i and j as,

$$G_w(i, j) = \#\{\hat{G}_b(i, j) = 1\}_{b=1}^{B}$$

Step 5: Further filtering tests the following hypothesis to select an edge in the G_w.
$H_0^{n_{ij}}$: The number of networks n_{ij} in the ensemble with an edge between process variables i and j is less than $n_0(\alpha)$ where α is the significance level. The n_{ij} follows a binomial distribution of $B(p, N)$ and

$$p_{ij} = Pr(n \geq n_{ij}) = \sum_{n=n_{ij}}^{B} \binom{B}{n} p^n (1 - p)^{B-n}$$

where p_{ij} is the probability that an edge is inferred by chance between process variables i and j more than n_{ij} times. The p is the probability that two process variables are connected randomly in the B network ensembles. The probability p that n_{ij} is connected by chance in an B network ensemble is estimated by computing the ratio of the total number of inferred edges in the ensemble and all possible edges as $B \times n(n-1)/2$, where n is the number of process variables.
Step 6: We test one hypothesis for each pair of process variables as discussed in step 4. Thus, we simultaneously test $n \times (n-1)/2$ hypotheses. In this multiple hypotheses testing case, we apply the multiple testing correction (MTC) and apply the Benjamini and Hochberg (FDR) [14] control in the analyses. The final network from G is defined as:

$$G(i, j) = \begin{cases} 1 \text{ if } p_{adjusted}(i, j) < \alpha \\ 0 \text{ otherwise} \end{cases}$$

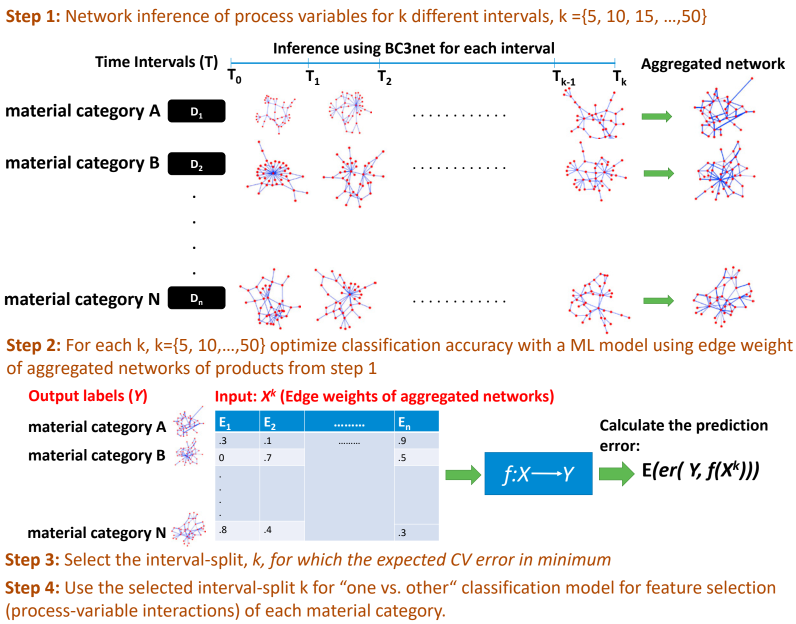

Fig. 1. Schematic representation of aggregated networks of process variables and optimizing the interval split from process variable data of different material categories.

3.4. Network aggregation for time intervals

For the constructed aggregated network over time intervals, we first split the process variable data $X(m_i, p_j)$ for product p_j on machine m_i into $k = \{5, 10, \ldots 50\}$ intervals. The production data are obtained for intervals defined as: $T = \{T_1, T_2, \ldots, T_k\}$ into $|T| = k$ intervals, where $T_i = [t_a, t_b]$, $T_{i+1} = [t_{b+1}, t_c]$, and $b - a = c - (b + 1)$. The network G_{T_k,P_i} is inferred and discussed in Section 3.3 for the process variable data $X(T_k, P_i)$ of the k^{th} interval and product P_i. We inferred networks for k intervals that are aggregated network ensembles from the k non-overlapping time intervals. The edge weights of the aggregated network are defined as:

$$G_{P_i}(i, j) = \frac{\sum_{k=1}^{|T|} I(G_{T_k,P_i}(i, j))}{|T|}$$

where $I()$ is an indicator function that is 1 if an edge exists between process variables i and j. We use edge weights as features for the ML model to identify key features (interactions) for different material categories.

3.5. GLM for product classification and feature selection

To develop GLMs that classify material categories, we first use the constructed aggregated network over time intervals G_{P_i} for each product P_i by splitting the data into several (k) time intervals. A schematic diagram of the method is shown in Figure 1. We then use the weights of the edges for G_{P_i} as input features and material categories as the response variables. Finally, the GLM with elastic net regularization is used for feature selection for the“one vs. other” two-class classification model for each material category. We select the non-zero coefficients as key features (interactions) for the material category classification. We apply the GLM with regularization twice. First, we apply the GLM for multi-class classification over 7 material categories to select the optimal interval split over $k = \{5, 10, 15, \dots 50\}$. We select an interval split that provides the minimum classification error. We select the interval split of $k = 20$ for which the multi-class classification error is minimum. We then apply the GLM for the “one vs. others” two-class classification for 8 material categories. For each GLM, we apply regularization (elastic net penalty) to fine-tune the model and enable feature selection. A brief algorithmic description is given in Algorithm 1.

Algorithm 1 Steps for predictability analysis of material categories

```
D = D_1, D_2, ..., D_n are |D| data-sets of y_L = {y_1, y_2, ... y_p} material categories.
Let Z_T(k) be a zero matrix with |D| rows and n × (n − 1)/2 columns representing pairs of process variables.
Let intrvl = {5, 10, 15, ... 50}
for t=1 to |intrvl| do
    k = intrvl[t]
    Let T(k) = {T_1, T_2, ... T_k} be |T| = k intervals.
    for i=1 to |D| do
      G_agg = []                                                        ▹ Network object array of size |T|
        for j=1 to |T| do
            G_{D_i,T_j} = BC3NET(D_i(T_j)))
            G_agg[j] = G_{D_i,T_j}
        end for
        G_{P_i} = aggregate(G_agg)
        Z_T(k)[i, ] = W(u, v), where (u, v) ∈ E(G_{P_i}) and W(u, v) is a weight vector of size n × (n − 1)/2.
    end for
    Let, y_L = f(Z_T(k)) be the classification model to optimize the fitting parameters and predict material categories y_L.
    R_k := Calculate log-loss to evaluate the predicted results.
end for
Select k from intrvl for which R_k is a minimum.
Use Z_T(k) for "one vs. other" classification of 8 material categories.
Select key features for each material category.
```

3.6. Selecting key features (interactions)

The key predictors (process variable interactions) of material categories shown in Table 6 are selected as follows.

- Let the GLM with regularization (elastic net penalty) select the process variable interaction Z_n as a key feature (non-zero coefficient) that classifies a material category P_i as a binary classification “one vs. others” model for data set $Z^b_{T(k)}$ where $\{Z^b_{T(k)}\}^B_{b=1}$ is a bootstrapped data set from $Z_{T(k)}$. The $Z_{T(k)}$ is the edge feature matrix of different products (rows) and edge weights (columns) as obtained by aggregating inferred networks for different intervals of the production data. If the average value of the selected predictor variable Z_n in the material category P_i is greater than the average value in the remaining categories ($\bar{Z}_n(P_i) > \bar{Z}_n(P_{others})$), then the importance of Z_n is calculated as:

$$w_{y,Z_n} = \frac{\#(z_n \text{ is selected as a key feature by the GLM})}{|B|}$$

- Select Z_n as a key feature that classifies y_{P_i} if $w_{y,Z_n} > \alpha$. Here, we select $\alpha = 0.5$.

4. Results

4.1. Interval selection

The first step of the analysis is to train the model using the GLM with regularization (elastic net penalty) for multi-class classification problems to optimize the best interval split using the algorithm from Algorithm 1. We then tune models with ten-fold cross-validation and up-sampling using various hyper-parameter combinations for different interval splits. The results of the best performing models classifying 8 material categories are shown in Table 2. The minimum classification error (log-loss) is found for $k = 20$. This indicates splitting the production data for all material categories into 20 time intervals optimizes the best classification with a minimum prediction error over the 8 material categories. The much lower prediction error also highlights that the process variable interactions are significantly different for various material categories and are likely to be similar within a single material category. The interval split $k = 20$ is further used for the "one vs. others" classification to identify key interactions over different material categories and to create and analyze the aggregated network.

Table 2. Classification accuracy of models when the process variable data are split into different intervals of size $k = \{5, 10, 15, \ldots 50\}$, and the input-features are the weighted edges of the aggregated network across time intervals.

Interval split	log-loss		Accuracy		Per class accuracy	
	Mean	S. D.	Mean	S. D.	Mean	S. D.
K=5	0.932	0.199	0.691	0.092	0.657	0.048
K=10	0.872	0.187	0.723	0.083	0.694	0.042
K=15	0.877	0.262	0.744	0.068	0.748	0.092
K=20	**0.798**	**0.236**	**0.749**	**0.088**	**0.730**	**0.099**
K=25	0.897	0.243	0.733	0.062	0.724	0.078
K=30	0.839	0.234	0.744	0.097	0.738	0.118
K=35	0.806	0.242	0.756	0.095	0.749	0.105
K=40	0.849	0.262	0.735	0.096	0.733	0.089
K=45	0.868	0.227	0.716	0.097	0.725	0.102
K=50	0.852	0.249	0.731	0.119	0.728	0.132

4.2. Aggregated network across all samples with steady interactions over time

We generate an aggregated network to show the overall interactions across the 215 products in 8 material categories that remain steady over time. We perform the following for aggregated networks across products that have steady interactions over $k = 20$ intervals. Select a $G^s_{P_i} \subset G_{P_i}$ for product P_i, where $\{w(e_i) \geq 0.95\}_{i=1}^{|E(G_{P_i})|}$, $e_i \in E(G_{P_i})$, and the final aggregated network is $G_{Aggregated} = G^s_{P_1} \cup G^s_{P_2}, \ldots G^s_{P_n}$. The resulting network is shown in Figure 2 and contains 39 process variables as vertices and 108 edges. We further apply the fast greedy algorithm to find communities (groups) of process variables in the aggregated network. The modularity as computed from the fast greedy algorithm is 0.451, which suggests a community structure that contains 6 modules. The list of process variables in different modules is shown in Table 3. Of note, the identified modules are mostly grouped into the same category. For example, in module 1, the *switch-over*-related process variables are grouped in the same module. In module 2, the *torque*-related process variables are grouped along with the process variables that measure the temperatures of different zones. The process variable *plasticizing number* is directly linked to the temperature zone and torque-related process variables.

In the module three, cycle-time-related process variables are grouped together. The other process variables in module 3 are *cooling time*, *pump pressure*, and *temperature zone 3*. The *cooling time* also shows interactions with the *torque mean value current cycle* and *plasticizing number* of module 2 and the *material cushion holding pressure* of module 4. The *pump pressure* has interaction with the *plasticizing number* process variable of module 2. Module 4 primarily contains the *material cushion, injection, and pressure*-related process variables and the *shot volume* and *plasticizing volume* process variables. The process variable in module 4 shows 12 interactions (edges) with the module 3 process variables. The module 5 process variables are the *closing force* and *clamping force peak value*, which do not show any interactions with other process variables. However, this may not always be true and only suggests these two variables are more steady compared to other interactions over time for a production process. Module 6 contains three *back pressure*-related process variables and *cycle time holding pressure*. The *back pressure peak value* shows interactions with the injection time and torque peak value from module 2. We further calculate the degree of process variables for the aggregated network and average degree of all process variables, which is the average degree calculated from the time-wise aggregated network of different product networks, as shown in Table 4. The maximum and minimum degrees are 11 for *cooling time* and 1 for *clamping force value-closing force switch-over position value* processing variables, respectively, in the aggregated network, which reflects steady interactions over time.

The maximum and minimum average degrees are 5.35 for the *cycle time automatic* and 1.10 for *switch-over volume value*, respectively. The calculated degree of the aggregated network and over all products are not the same and can

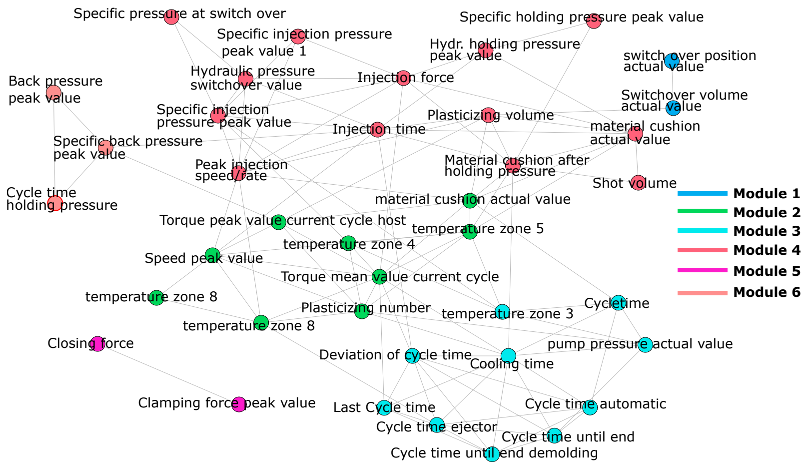

Fig. 2. Aggregated network with edges that are steady over time with the interval split K=20. The 6 identified modules are shown in different colors.

vary significantly. The main purpose of showing these results is to identify sensitive process variables that can be affected or affect other process variables in production processes. The interactions indicate that such links cannot be the same over all production processes and can be machine-, material-, and product-specific. The other case of low-degree nodes indicates that such process variables are less sensitive and not affected by other process variables nor do they affect others. However, a domain expert's interpretation is required to validate the correlation between the sensitivity of the process variables and the estimated degrees of process variables in the network. Sensitivity plays a key role in controlling other process variables or those to be controlled by others and can be understood using node importance measures for material-specific inferred networks. However, evaluating node-importance measures with the process variable sensitivity is outside the scope of this paper.

Table 3. Process variables in different modules of aggregated network across products for best interval split (k=20).

S. No.	Module	Process variables
1	Module 1	Switch-over position actual value, switch-over volume actual value
2	Module 2	Torque peak value current cycle host, torque mean value current cycle, speed peak value, plasticizing number, temperature zone 1, temperature zone 4, temperature zone 5, temperature zone 8
3	Module 3	Last Cycle time, cooling time, deviation of cycle time, pump pressure actual value, temperature zone 3, cycle time until end, cycle time, cycle time ejector, cycle time until end demolding, cycle time automatic
4	Module 4	Plasticizing volume, peak injection speed/rate, material cushion actual value, material cushion actual value, material cushion actual length, material cushion after holding pressure, injection time, injection force, specific injection pressure peak value, hydraulic pressure switch-over value, hydr. holding pressure peak value, specific injection pressure peak value 1, specific pressure at switch-over, specific holding pressure peak value, shot volume
5	Module 5	Closing force, clamping force peak value
6	Module 6	Specific back pressure peak value, back pressure peak value, cycle time holding pressure

4.3. material category-specific key interaction features

The feature selection (key process variable interactions) of each material category label of a material category is "1" and the remaining categories are "0" as in "one vs. other." The binary labels to model the two-class classification of each material category are used as the response variable for the classification model. A specific material category is labeled as "1" in each model, and all remaining material categories are labeled as "others." We apply the GLM with regularization, up-sampling, and 10-fold cross-validation for each material category classification. The evaluation metrics for classifications of different products are shown in Table 5. All classification models show strong accuracy when classifying different products in the two-class classification model as indicated by the average log-loss, average AUC, and other metrics. For the feature selection, we apply the method discussed in Section 3.6. The resulting key features (process variable interactions) are shown in Table 6 and as networks in Figure 3. The features are ordered by their importance for each material category. The interactions between two process variables are identified as key interactions and are separated by " – ", e.g. ($a - b$: a interact with b). Figure 3 shows the visualization of connecting

Table 4. Degrees of process variables in the aggregated network and average degree by averaging across products.

Process variables	degree in aggregated network	average degree
Cooling time	11.00	2.44
Torque mean value current cycle	11.00	3.03
Injection time	10.00	4.60
Material cushion after holding pressure	9.00	1.60
Peak injection speed/rate	9.00	2.29
Specific injection pressure peak value	8.00	4.13
Deviation of cycle time	8.00	3.00
temperature zone 5	8.00	4.30
Plasticizing number	8.00	4.52
Injection force	8.00	2.60
Speed peak value	8.00	4.84
Material cushion actual value material cushion actual length	8.00	2.20
Material cushion actual value	8.00	3.33
Cycle time ejector	7.00	4.53
Plasticizing volume	7.00	3.14
Cycle time automatic	6.00	5.35
Cycle time until end demolding	6.00	2.73
Cycle time until end	6.00	3.15
Temperature zone 1	6.00	4.29
Torque peak value current cycle host	6.00	4.57
Hydraulic pressure switch-over value	5.00	3.91
Cycle time	5.00	4.21
Temperature zone 3	5.00	2.18
Pump pressure actual value	5.00	3.04
Temperature zone 4	5.00	2.53
Last Cycle time	5.00	1.95
Hydr. holding pressure peak value	4.00	3.66
Specific back pressure peak value	4.00	4.49
Specific injection pressure peak value 1	3.00	4.00
Cycle time holding pressure	2.00	2.66
Back pressure peak value	2.00	2.47
Shot volume	2.00	2.36
Specific holding pressure peak value	2.00	2.57
Specific pressure at switch-over	2.00	2.91
Temperature zone 8	2.00	4.40
Switch-over volume actual value	2.00	1.10
Clamping force peak value	1.00	2.27
Closing force	1.00	1.82
Switch-over position actual value	1.00	2.97

variables in an understandable manner. Although the graphs are not connected, they show that some interactions are independent and with limited impact. They also depend on the threshold α for filtering key features. In Figure 3, different material categories indicate the process variables for key interactions overlap and join into larger connected components. For example, the material category PA shows two large connecting components that contain more than two process variables at 11 and 7. The material category *PE* shows the largest connecting component of 14 process variables. The key interactions that join different process variables indicate the complexity of the interactions, which must be tuned collectively to ensure the efficient production of a specific product. The key features of the 8 material

Table 5. Evaluation measures of classification models for material categories (one vs. others) when the process variable data is split into $k = 20$ different intervals, and the input features are the weighted edges of the inferred networks at different intervals and aggregated networks across time intervals.

	Log-loss		AUC		Precision		Recall		Accuracy	
material category	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
PA	0.303	0.095	0.937	0.053	0.910	0.062	0.989	0.024	0.917	0.047
PC	0.178	0.106	0.933	0.094	0.972	0.033	1.000	0.000	0.973	0.032
ABS	0.301	0.129	0.833	0.165	0.939	0.041	0.981	0.042	0.932	0.048
PP	0.201	0.134	0.968	0.031	0.943	0.063	1.000	0.000	0.957	0.049
POM	0.144	0.180	0.890	0.314	0.986	0.022	1.000	0.000	0.987	0.022
PS	0.073	0.138	1.000	0.000	1.000	0.000	1.000	0.000	1.000	0.000
PE	0.175	0.107	0.823	0.283	0.976	0.027	1.000	0.000	0.977	0.025
PET	0.002	0.003	1.000	0.000	1.000	0.000	1.000	0.000	1.000	0.000

categories are the common process variables. For example, the process variable *injection time* interacts with other process variables (interaction) in 13 different (maximum) features, which shows its importance as one of the key process variables during production. Similarly, the other process variables of *cooling time*, *cycle time*, *plasticizing* *volume*, *peak injection speed/rate*, *material cushion actual value*, *specific injection pressure peak value torque mean current value cycle*, *speed peak value*, and *temperature zone 5* interact with more than 6 other process variables. There are some key features (interactions) that are important in more than 1 material category. For example, the interaction between the *torque peak value current cycle* and *plasticizing number* is a key feature in 4 material categories. We further calculate feature overlap, as shown in Table 7. The overlap indicates that a small fraction of features are common between material categories, which suggests that many key features are unique to their respective material categories.

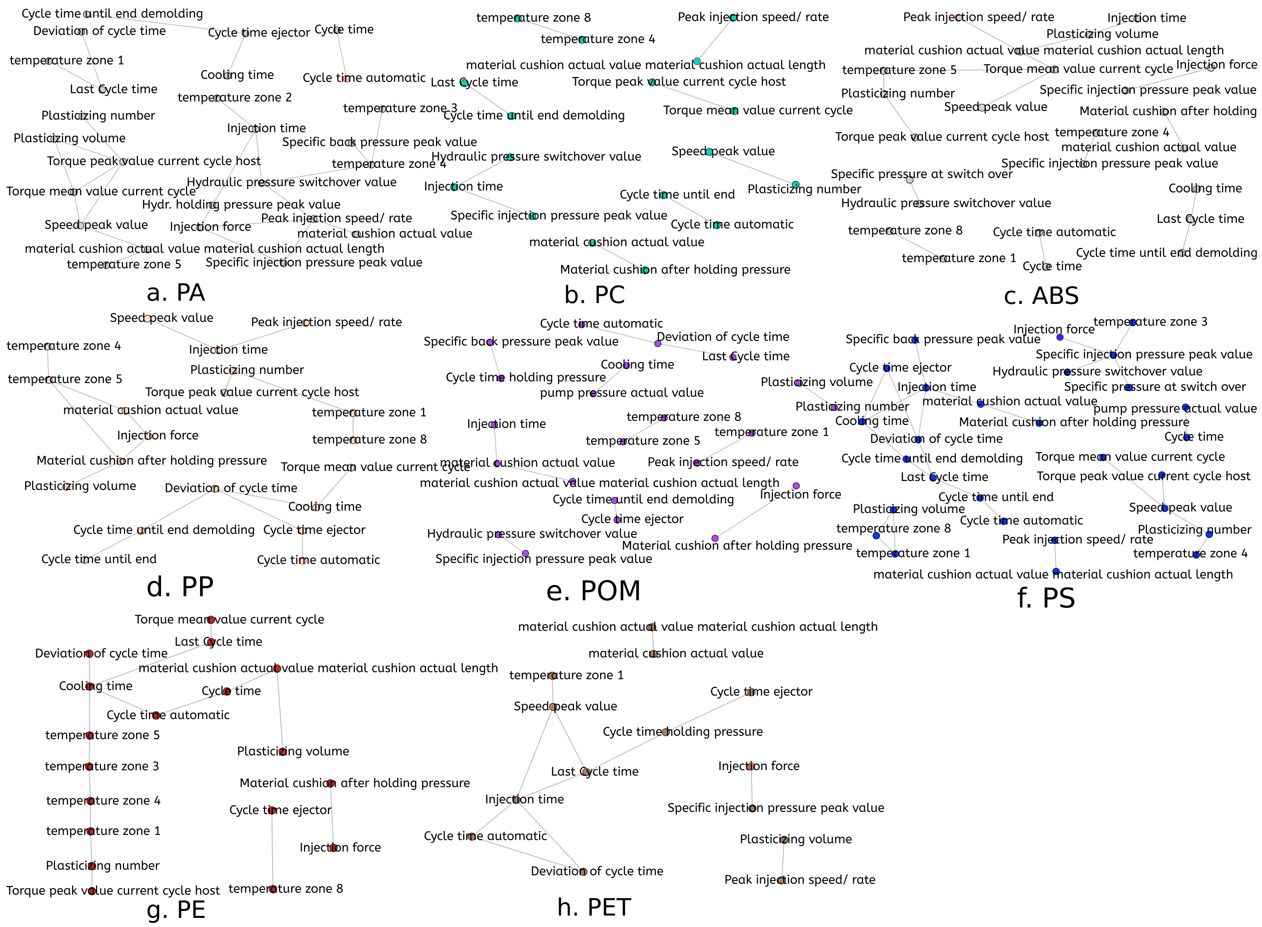

Fig. 3. Key interaction features of each material category shown as a network

4.4. Comparison with related studies

We compare 10 process variables used to predict the product quality and scrap rate in previous studies. The comparison in Table 8 shows 10 key process variables and corresponding variables that interact with them. We see that all key process variables interact with other process variables in at least 4 material categories. However, their interactions are not the same in different material categories. This suggests that we should consider other interacting process variables found in specific material categories and key process variables for quality optimization.

Table 6. Key process variable interactions as separated by different material categories. These features are identified by applying the GLM with regularization.

material categories	Total features	Process variable interactions
PA	25	Cycle time – Cycle time automatic, Torque peak value current cycle host – Speed peak value, Specific injection pressure peak value – Peak injection speed/rate, Cycle time ejector – Cycle time until end demolding, Torque mean value current cycle – Speed peak value, Injection force – Hydr. holding pressure peak value, Peak injection speed/rate – material cushion actual value, Injection force – Peak injection speed/rate, Torque mean value current cycle – Torque peak value current cycle host, Injection force – Specific injection pressure peak value, Hydraulic pressure switchover value – Injection time, Torque peak value current cycle host – Plasticizing number, Peak injection speed/rate – Hydraulic pressure switchover value, Plasticizing volume – Torque peak value current cycle host, Plasticizing volume – Torque mean value current cycle, Last Cycle time – Deviation of cycle time, Last Cycle time – Temperature zone 1, Temperature zone 3 – Temperature zone 4, Hydraulic pressure switchover value – Temperature zone 4, Cooling time – Cycle time ejector, Injection time – Temperature zone 2, Hydr. holding pressure peak value – Injection time, Specific back pressure peak value – Temperature zone 4, Material cushion actual value material cushion actual length – Temperature zone 5, Speed peak value – Material cushion actual value material cushion actual length
PC	9	Torque mean value current cycle – Torque peak value current cycle host, Peak injection speed/rate – material cushion actual value material cushion actual length, Specific injection pressure peak value – Injection time, material cushion actual value – Material cushion after holding pressure, Last Cycle time – Cycle time until end demolding, Hydraulic pressure switchover value – Injection time, Temperature zone 4 – Temperature zone 8, Speed peak value – Plasticizing number, Cycle time until end – Cycle time automatic
ABS	16	Torque peak value current cycle host – Plasticizing number, Plasticizing volume – Injection time, Temperature zone 1 – Temperature zone 8, Material cushion actual value – Material cushion after holding pressure, Plasticizing number – Temperature zone 5, Injection force – Specific injection pressure peak value 1, Torque mean value current cycle – Speed peak value, Torque mean value current cycle – Temperature zone 5, Plasticizing volume – material cushion actual value material cushion actual length, Peak injection speed/rate – Material cushion actual value material cushion actual length, Cycle time – Cycle time automatic, Last Cycle time – Cycle time until end demolding, Torque mean value current cycle – Material cushion actual value material cushion actual length, Last Cycle time – Cooling time, Specific injection pressure peak value – temperature zone 4, Hydraulic pressure switchover value – Specific pressure at switch-over
PP	17	Injection force – Material cushion actual value, Deviation of cycle time – Cycle time until end demolding, Temperature zone 4 – Temperature zone 5, Deviation of cycle time – Cycle time ejector, Temperature zone 1 – Temperature zone 8, Peak injection speed/rate – Injection time, Plasticizing volume – Material cushion after holding pressure, Cycle time ejector – Cycle time automatic, Plasticizing number – Temperature zone 1, Torque mean value current cycle – Cooling time, Injection force – Material cushion after holding pressure, Cooling time – Deviation of cycle time, Cycle time until end – Cycle time until end demolding, Material cushion after holding pressure – Temperature zone 5, Material cushion actual value – Temperature zone 5, Torque peak value current cycle host – Plasticizing number, Speed peak value – Injection time
POM	25	Injection force – Specific injection pressure peak value, Specific injection pressure peak value – Hydraulic pressure switchover value, Cycle time until end – Cycle time automatic, Cooling time – Cycle time until end demolding, Plasticizing number – Temperature zone 4, Torque mean value current cycle – Speed peak value, Specific injection pressure peak value – Temperature zone 3, Specific injection pressure peak value – Specific pressure at switch-over, material cushion actual value – Material cushion after holding pressure, Peak injection speed/rate – Material cushion actual value material cushion actual length, Last Cycle time – Cycle time until end demolding, Torque peak value current cycle host – Speed peak value, Plasticizing volume – temperature zone 1, Material cushion actual value – Injection time, Injection time – Cooling time, Pump pressure actual value – Cycle time, Speed peak value – Plasticizing number, Temperature zone 1 – Temperature zone 8, Plasticizing volume – Temperature zone 8, Injection time – Deviation of cycle time, Deviation of cycle time – Cycle time ejector, Specific back pressure peak value – Injection time, Last Cycle time – Cycle time until end, Cooling time – Cycle time ejector, Last Cycle time – Deviation of cycle time
PS	12	Material cushion actual value – Material cushion actual value material cushion actual length, Specific injection pressure peak value – Hydraulic pressure switch-over value, Deviation of cycle time – Cycle time automatic, Specific back pressure peak value – Cycle time holding pressure, Last Cycle time – Deviation of cycle time, pump pressure actual value – Cooling time, Cycle time ejector – Cycle time until end demolding, material cushion actual value – Injection time, Plasticizing volume – Plasticizing number, Injection force – Material cushion after holding pressure, Temperature zone 5 – Temperature zone 8, Peak injection speed/rate – Temperature zone 1
PE	15	Injection force – Material cushion after holding pressure, Material cushion actual value material cushion actual length – Cycle time, Plasticizing volume – Material cushion actual value material cushion actual length, Plasticizing number – Temperature zone 1, Torque peak value current cycle host – Plasticizing number, Cooling time – Cycle time automatic, Cycle time – Cycle time automatic, Torque mean value current cycle – Last Cycle time, Temperature zone 5 – Cooling time, Temperature zone 1 – Temperature zone 4, Temperature zone 8 – Cycle time ejector, Temperature zone 3 – Temperature zone 5, Last Cycle time – Cooling time, Cooling time – Deviation of cycle time, Temperature zone 3 – Temperature zone 4
PET	12	Injection time – Deviation of cycle time, Deviation of cycle time – Cycle time automatic, Plasticizing volume – Peak injection speed/rate, material cushion actual value – Material cushion actual value material cushion actual length, Speed peak value – Last Cycle time, Injection force – Specific injection pressure peak value, Last Cycle time – Injection time, Speed peak value – Injection time, Last Cycle time – Cycle time holding pressure, Speed peak value – temperature zone 1, Cycle time ejector – Cycle time holding pressure, Injection time – Cycle time automatic

Table 7. Proportion of overlapped features in different material categories.

	PA	PC	ABS	PP	POM	PS	PE	PET
PA	1.000	0.062	0.079	0.024	0.111	0.057	0.081	0.028
PC	0.062	1.000	0.136	0.000	0.172	0.000	0.000	0.000
ABS	0.079	0.136	1.000	0.065	0.139	0.000	0.148	0.000
PP	0.024	0.000	0.065	1.000	0.050	0.036	0.143	0.036
POM	0.111	0.172	0.139	0.050	1.000	0.088	0.000	0.057
PS	0.057	0.000	0.000	0.036	0.088	1.000	0.038	0.091
PE	0.081	0.000	0.148	0.143	0.000	0.038	1.000	0.000
PET	0.028	0.000	0.000	0.036	0.057	0.091	0.000	1.000

Table 8. Process variables from previous studies that play a key role in product quality and their corresponding interacting process variables.

	Process variables from previous studies which play a key role in product quality	interaction with other process variables	material category
1	Material cushion after holding pressure	Material cushion actual value, temperature zone 5, plasticizing volume, injection force	PC, ABS, PP, POM, PS, PE
2	Material cushion actual value	Peak injection speed/rate, Material cushion after holding pressure, temperature zone 5, injection force, injection time, material cushion actual value material cushion actual length	PA, PC, ABS, PP, POM, PS, PET
3	Specific injection pressure peak value	Peak injection speed/rate, Injection force, Injection time, temperature zone 4, hydraulic pressure switch-over value, temperature zone 3, specific pressure at switch-over	PA, PC, ABS, POM, PS, PET
4	Hydraulic pressure switch-over value	Injection time, temperature zone 4, Peak injection speed/rate, Specific pressure at switch-over, Specific injection pressure peak value	PA, PC, ABS, POM, PS
5	Injection force	Hydr. holding pressure peak value, peak injection speed/rate, specific injection pressure peak value, Specific injection pressure peak value 1, material cushion actual value, material cushion after holding pressure	PA, ABS, PP, POM, PS, PE, PET
6	Torque mean value current cycle	Speed peak value, torque peak value current cycle host, plasticizing volume, temperature zone 5, material cushion actual value material cushion actual length, cooling time, last cycle time	PA, PC, ABS, PP, POM, PE
7	Cycle time	Cycle time automatic, pump pressure actual value, material cushion actual value material cushion actual length	PA, ABS, POM, PE
8	Cooling time	Cycle time ejector, last cycle time, deviation of cycle time, torque mean value current cycle, cycle time until end demolding, injection time, pump pressure actual value, cycle time automatic, temperature zone 5	PA, ABS, PP, POM, PS, PE
9	Injection time	Temperature zone 2, Hydraulic pressure switch-over value, Hydr. holding pressure peak value, specific injection pressure peak value, plasticizing volume, peak injection speed/rate, speed peak value, cooling time, deviation of cycle time, material cushion actual value, specific backpressure peak value, cycle time automatic, last cycle time	PA, PC, ABS, PP, POM, PS, PET
10	Peak injection speed/rate	Material cushion actual value, hydraulic pressure switch-over value, specific injection pressure peak value, injection force, material cushion actual value material cushion actual length, injection time, temperature zone 1, plasticizing volume	PA, PC, ABS, PP, POM, PS, PET

5. Conclusion

This paper analyzes process variable data for injection molding to identify interdependent process variables as key features in different material categories. We provide a systematic analysis of process variable data during the production of different material categories to understand key interactions between them that play key roles in quality production. We infer networks of process variables for the production data of each product as grouped into 8 material categories. The applied BC3net has the primary advantage that it is a network ensemble approach that uses mutual information to calculate linear and non-linear relationships between process variables. The other advantage is that the bootstrapped approach does not require any assumptions about the data distribution. Various parameters are set during the production process that can be tuned or changed in-between production processes. Therefore, static correlation or association measure-based models generate noise or false positives and ignore changes in variable interactions over time. As a result, we use time-wise aggregation of the network to calculate the stability and variability of interactions. Another critical point is that the number of time intervals (interval size) impacts the classification of material categories, which means that the inferred interactions are also affected by the interval size. The underlying key learning objective from the analysis here is that the modular nature of process variable interactions that are steady over time and those that vary over time as a network differ significantly for various material categories. The distinct interaction patterns indicate the complexity of process parameter selection for optimization to ensure product quality and efficient production for different product types.

Domain experts can further interpret these analysis results to classify the most important process parameters and the corresponding interactions in production processes for specific material categories considering various other external factors. The structure of process variables interactions for each material category and their modules in aggregated networks can be further utilized to develop functional descriptors that feature process fingerprints. Such fingerprints are described as direct or indirect quality indicators to be monitored or controlled for efficient production [13, 29]. We can utilize and evaluate the key interactions of each material category as estimated from the analysis between different process variables to predict future effects in advance so that required actions can be initiated by readjusting the process parameters and other relevant maintenance steps. Additionally, previous understandings of the role of process variables in efficient and quality production can be compared with current interactions shown as aggregated networks and material-specific interactions from various material categories. The results from the analysis provide a better understanding of material category-specific process variable dependencies over time, which can be utilized to update the process meta-knowledge for production planning and production process calibration.

Some of the limitations in the current analysis are the uneven sample sizes in different material categories and the exclusion of 18 process variables that are present in only a few data sets. For future work, we would like to acquire production data for different material categories that include detailed information of the produced scrap, material-specific characteristics, material-specific details, and external effect information such as machine-age, surrounding temperature, and humidity to develop a robust and explainable model that could identify material category-specific key features that play crucial roles in maintaining product quality.

ACKNOWLEDGMENTS

This paper is a part of X-pro project. The project is financed by research subsidies granted by the government of UpperAustria.

References

[1] Gökmen Altay and Frank Emmert-Streib. Structural influence of gene networks on their inference: analysis of c3net. *Biology Direct*, 6(1):1–16, 2011.

[2] Kestutis Baltakys, Juho Kanniainen, and Frank Emmert-Streib. Multilayer aggregation with statistical validation: Application to investor networks. *Scientific reports*, 8(1):1–12, 2018.

[3] Mahmoud A. Barghash and Faisal Alkhannan Alkaabneh. Shrinkage and warpage detailed analysis and optimization for the injection molding process using multistage experimental design. *Quality Engineering*, 26(3):319–334, 2014.

[4] M. Berins and Society of the Plastics Industry. *Plastics Engineering Handbook Of The Society Of The Plastics Industry*. Plastics Engineering. Springer US, 1991.

[5] Tao C. Chang and Ernest Faison III. Shrinkage behavior and optimization of injection molded parts studied by the taguchi method. *Polymer Engineering & Science*, 41(5):703–710, 2001.

[6] Wen-Chin Chen, Manh-Hung Nguyen, Wen-Hsin Chiu, Te-Ning Chen, and Pei-Hao Tai. Optimization of the plastic injection molding process using the taguchi method, rsm, and hybrid ga-pso. *The International Journal of Advanced Manufacturing Technology*, 83(9):1873–1886, Apr 2016.

[7] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[8] Xuan-Phuong Dang. General frameworks for optimization of plastic injection molding process parameters. *Simulation Modelling Practice and Theory*, 41:15 – 27, 2014.

[9] Ricardo de Matos Simoes and Frank Emmert-Streib. Bagging statistical network inference from large-scale gene expression data. *PloS one*, 7(3):e33624, 2012.

[10] S. Ebnesajjad and P.R. Khaladkar. *Fluoropolymer Applications in the Chemical Processing Industries: The Definitive User's Guide and Databook*. Plastics Design Library. Elsevier Science, 2004.

[11] Frank Emmert-Streib and Matthias Dehmer. High-dimensional lasso-based computational regression models: regularization, shrinkage, and selection. *Machine Learning and Knowledge Extraction*, 1(1):359–383, 2019.

[12] Célio Fernandes, António José Pontes, Júlio César Viana, and António Gaspar-Cunha. Modeling and optimization of the injection-molding process: A review. *Advances in Polymer Technology*, 37(2):429–449, 2018.

[13] Nikolaos Giannekas, Per Magnus Kristiansen, Yang Zhang, and Guido Tosello. Investigation of product and process fingerprints for fast quality assurance in injection molding of micro-structured components. *Micromachines*, 9(12):661, 2018.

[14] Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.

[15] Suchana A. Jahan and Hazim El-Mounayri. Optimal conformal cooling channels in 3d printed dies for plastic injection molding. *Procedia Manufacturing*, 5:888 – 900, 2016. 44th North American Manufacturing Research Conference, NAMRC 44, June 27-July 1, 2016, Blacksburg, Virginia, United States.

[16] Satadru Kashyap and Dilip Datta. Process parameter optimization of plastic injection molding: a review. *International Journal of Plastics Technology*, 19(1):1–18, Jun 2015.

[17] Yongfeng Li, Fei Tao, Ying Cheng, Xianzhi Zhang, and A.Y.C. Nee. Complex networks in advanced manufacturing systems. *Journal of Manufacturing Systems*, 43:409 – 421, 2017. Special Issue on the 12th International Conference on Frontiers of Design and Manufacturing.

[18] Ting Mao, Yun Zhang, Yufei Ruan, Huang Gao, Huamin Zhou, and Dequn Li. Feature learning and process monitoring of injection molding using convolution-deconvolution auto encoders. *Computers & Chemical Engineering*, 118:77 – 90, 2018.

[19] Duoc T Nguyen. Injection molding scrap reduction: a study in the relationships of plastics processing methods. 2004.

[20] Erfan Oliaei, Behzad Shiroud Heidari, Seyed Mohammad Davachi, Mozhgan Bahrami, Saeed Davoodi, Iman Hejazi, and Javad Seyfi. Warpage and shrinkage optimization of injection-molded plastic spoon parts for biodegradable polymers using taguchi, anova and artificial neural network methods. *Journal of Materials Science & Technology*, 32(8):710 – 720, 2016.

[21] Catharina Olsen, Patrick E Meyer, and Gianluca Bontempi. On the impact of entropy estimation on transcriptional regulatory network inference based on mutual information. *EURASIP Journal on Bioinformatics and Systems Biology*, 2009:1–9, 2008.

[22] Michael Packianather, Christian Griffiths, and Wan Kadir. Micro injection moulding process parameter tuning. *Procedia CIRP*, 33:400 – 405, 2015. 9th CIRP Conference on Intelligent Computation in Manufacturing Engineering - CIRP ICME '14.

[23] Wei Qin, Dongye Zha, and Jie Zhang. An effective approach for causal variables analysis in diesel engine production by using mutual information and network deconvolution. *Journal of Intelligent Manufacturing*, pages 1–11, 2018.

[24] B. Ribeiro. Support vector machines for quality monitoring in a plastic injection molding process. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 35(3):401–410, Aug 2005.

[25] D.V. Rosato and M.G. Rosato. *Injection Molding Handbook*. Springer US, 2012.

[26] Gurjeet Singh and Ajay Verma. A brief review on injection moulding manufacturing process. *Materials Today: Proceedings*, 4(2, Part A):1423 – 1433, 2017. 5th International Conference of Materials Processing and Characterization (ICMPC 2016).

[27] G. Taguchi, E.A. Elsayed, and T.C. Hsiang. *Quality engineering in production systems*. McGraw-Hill series in industrial engineering and management science. McGraw-Hill, 1989.

[28] G. Taguchi, S. Konishi, S. Konishi, and American Supplier Institute. *Orthogonal Arrays and Linear Graphs: Tools for Quality Engineering*. Taguchi methods. American Supplier Institute, 1987.

[29] Guido Tosello. Product/process fingerprint in micro manufacturing. *Micromachines*, 10:340, 05 2019.

[30] Shailesh Tripathi, Christian Mittermayr, David Muhr, and Herbert Jodlbauer. Large scale predictability analysis of process variables from injection molding machines. *Procedia Computer Science*, 180:545–560, 2021.

[31] Shailesh Tripathi., Sonja Strasser., Christian Mittermayr., Matthias Dehmer., and Herbert Jodlbauer. Approaches to identify relevant process variables in injection moulding using beta regression and svm. In *Proceedings of the 8th International Conference on Data Science, Technology and Applications - Volume 1: DATA,*, pages 233–242. INSTICC, SciTePress, 2019.

[32] Resit Unal and Edwin B. Dean. Taguchi approach to design optimization for quality and cost : An overview. 1991.

[33] Lixiang Wang, Wei Dai, Dongmei Sun, and Yu Zhao. Risk evaluation for a manufacturing process based on a directed weighted network. *Entropy*, 22(6):699, 2020.

[34] Fei Yin, Huajie Mao, and Lin Hua. A hybrid of back propagation neural network and genetic algorithm for optimization of injection molding process parameters. *Materials & Design*, 32(6):3457 – 3464, 2011.

[35] Fei Yin, Huajie Mao, Lin Hua, Wei Guo, and Maosheng Shu. Back propagation neural network modeling for warpage prediction and optimization of plastic products during injection molding. *Materials & Design*, 32(4):1844 – 1850, 2011.

[36] Boštjan Zafošnik, Uroš Božič, and Blaž Florjanič. Modelling of an analytical equation for predicting maximum stress in an injections moulded undercut geometry during ejection. *International Journal of Precision Engineering and Manufacturing*, 16(12):2499–2507, nov 2015.

[37] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, 67(2):301–320, 2005.