Approaches to Identify Relevant Process Variables in Injection Moulding using Beta Regression and SVM

Shailesh Tripathi[1], Sonja Strasser[1], Christian Mittermayr[2], Matthias Dehmer[1] and Herbert Jodlbauer[1]
[1]Production and Operations Management, University of Applied Sciences Upper Austria, Wehrgrabengasse 1-3, Steyr, Austria
[2]Greiner Packaging International GmbH, Greinerstrasse 70, 4550 Kremsmuenster, Austria

Keywords: Injection Moulding, Beta Regression, SVM, Scrap Rate Prediction.

Abstract: In this paper, we analyze data from an injection moulding process to identify key process variables which influence the quality of the production output. The available data from the injection moulding machines provide information about the run-time, setup parameters of the machines and the measurements of different process variables through sensors. Additionally, we have data about the total output produced and the number of scrap parts. In the first step of the analysis, we preprocessed the data by combining the different sets of data for a whole process. Then we extracted different features, which we used as input variables for modeling the scrap rate. For the predictive modeling, we employed three different models, beta regression with the backward selection, beta boosting with regularization and SVM regression with the radial kernel. All these models provide a set of common key features which affect the scrap rates.

1 INTRODUCTION

Injection moulding is regarded as the most important process to produce all kind of plastic products. Simply put, a melted polymer is injected into a mold cavity, packed under pressure and cooled until it has solidified enough. This is performed by an injection molding machine using an appropriate injection mold. During the whole process, the material, the mold design and the processing parameters of the injection molding machine interact with each other and determine the quality of the plastic product (Chang and Faison III, 2001). Since there is a huge variety of different processing parameters, the complexity of the process creates a very high effort to keep the quality characteristics under control. If the necessary quality characteristics cannot be achieved, the parts are discarded as scrap. Quality problems can be of different types, such as shrinkage, warpage, color and burn marks, surface texture quality, shape distortion, and other aesthetic defects (Kashyap and Datta, 2015). In a real world industrial production the scrap rate varies in different proportions during the production process. Variation in scrap rate depends on many process variables which are machine specific, product specific and material specific. The main objective of the quality control is to minimize the scrap rate during a production process. The scraps rate can be higher due to many uncontrolled process variables or their combinations. Most of the studies are based on controlling fewer variables through an experimental design and then estimating the quality of the products. However, in our analysis, we analyze data from the real-world injection molding process, which is recorded during the production process through different sensors which consist of > 75 process variables. The response variable is the proportion of the scraps which are produced during the injection molding process. The input variables are the statistical features of different process variables recorded during the production process. In this analysis, we do not distinguish between different types of scrap.

The main objective of the paper is to identify key features of process variables which affect the product quality for the purpose of monitoring in future to control the production quality of different products at different machines.

This paper contributes to the application of machine learning methods to identify common key process variables, which affect the quality of the production output represented by the scrap rate. The Structure of the paper is as follows: In Section 2 we provide a brief overview of the related work. In Section 3 we describe the details about the methods for

Tripathi, S., Strasser, S., Mittermayr, C., Dehmer, M. and Jodlbauer, H.
Approaches to Identify Relevant Process Variables in Injection Moulding using Beta Regression and SVM.
DOI: 10.5220/0007926502330242
In *Proceedings of the 8th International Conference on Data Science, Technology and Applications (DATA 2019)*, pages 233-242
ISBN: 978-989-758-377-3

the data preprocessing, filtering and normalization. In Section 4, we present details about the modeling and variable selection. In Section 5, we evaluate the proposed methods on recorded data and compare the results. Additionally, we compare our approach with two other methods (Ribeiro, 2005; Mao et al., 2018) which utilize SVM and a deep-learning approach to classify the product quality into different categories. In the final Section 6 we provide the concluding remarks about the results and our analysis.

2 LITERATURE REVIEW

A large number of studies have been performed for the quality optimization of the injection molding process. Many studies of quality optimization are based on Taguchi experimentation with a fewer number of key process variables, which are responsible for the product quality (Taguchi et al., 1987; Taguchi et al., 1989; Unal and Dean, 1991; Chang and Faison III, 2001; Barghash and Alkaabneh, 2014; Packianather et al., 2015; Chen et al., 2016; Oliaei et al., 2016; Jahan and El-Mounayri, 2016). Several computational approaches have been studied to optimize product quality. These computational techniques are based on gradient-based approaches, evolutionary algorithms and mixed approaches utilizing gradient-based approaches with evolutionary algorithms (Yin et al., 2011b; Zafošnik et al., 2015; Oliaei et al., 2016; Yin et al., 2011a; Chen et al., 2016). Reviews of the frameworks for the optimization of injection moulding methods are described by (Kashyap and Datta, 2015; Dang, 2014; Singh and Verma, 2017; Fernandes et al., 2018).

3 METHODOLOGY

In this section, we provide a brief overview of the methods we used for preprocessing, feature extraction, and the regression models. First, we describe the available data for our analysis and the major preprocessing steps. After splitting the data into different production lots (segments), we extract relevant features for the subsequent regression models. The aim is to train prediction models for the scrap rates to get information about the various setup parameters and process variables which have the highest impact on the scrap rates.

3.1 Data Collection

The data collection process starts by collecting raw data from 33 injection molding machines which are recorded into different files during a production process. Additionally, we use data from the enterprise resource planning (ERP) system. The relevant files for our analysis are the production files, the process files and an export file from the ERP system, which have the following information:

Production file: provides time stamps, cycle counter, tool-name, raw-material information, cavities and set cycle time.

Process data: provides time stamps, cycle counter, set cycle time and > 75 different process variables such as actual cycle time, temperature, pressure, volumes, positions, rotational speed, etc.

ERP data: provides the order number, material number, number of produced parts and number of scrap parts

The production and process data files contain data from a certain time period where multiple product types are produced. However, in the process data itself, there is no information about the product types. So we used the production data and the ERP data to split the process data into different segments according to the product type and order number these different segments based on product types and order numbers are described as process segments. Each process segment contains information about > 75 different process variables for the production process of a product type between the start time and the end time of production order. So the segment data file is a multivariate time series data file, where columns are the process variables and the rows are their respective measures at different time points.

Relevant meta information about each production order is collected in a separate master data file, including start and end time of production, order number, machine number, raw material number, total production output in units and the number of scrap parts in units.

3.2 Feature Extraction

For our analysis we extracted 6 statistical features of each process variable due to following reasons:

– The values recorded for different process variables are not recorded for fixed time intervals and not all process variables are recorded at the same time stamp.

– For the predictive modelling we want to make our results interpretable for the machine operators so that they can tune the key process variables appropriately.

Let us assume that X is a process variable which have values $X = \{x_{t_1}, x_{t_2}, \ldots x_{t_n}\}$ between time point t_1 and time point t_n. We extract common statistical features: mean, standard deviation, maximum, minimum, $M_1(X)$ and $M_2(X)$ in each process segment data file for each process variable. These statistical features are described in Appendix section 6.1. Thus for each process variables, we have six features. We store all the extracted features for each process segment data file in a data set, D.

3.3 Scrap Rate

The extracted features are the predictor variables for the prediction of the scrap rate. We add the scrap rate as the response variable in the data file D and calculate it, for each process segment p_i, from the masterdata file we extract the scraps samples and scaled the value of the scrap between 0 and 1, $r_{p_i} \in (0,1)$. The r_{p_i} is the new rescaled scrap rate for each process segment.

3.4 Data Filtering

For our analysis, we filter out those process segments which are running longer than 120 hours and shorter than 20 minutes. This information is extracted from the master data file. After this filtering step, D contains 1996 observations and 558 features.

Since we collected data from different injection molding machines, we do not have the same set of process variables for each machine and therefore deviating features for the process segments. In order to find a common set of features, we analyzed the frequency of observations of the extracted features. The result is shown in Figure 1. Then we select those features which appear in common at least in ≥ 600 observations. Thus we obtained 639 common observations for 327 features.

In the next step of filtering, we filter out highly correlated variables which have a high correlation of $\rho > .95$ with any of the other feature variables. Our approach to filter out such variables is described in Algorithm 3 in Appendix section 6.1. After this step, we have a $|L| = 66$ features. The correlation between these different features is shown in Figure 2

3.5 Data Normalization

We normalized each feature independently as follows: Suppose a feature $V_i = \{v_1, v_2, v_3, \ldots, v_n\}$ has n samples. We convert samples of each features into standard scores as follows:

$$V_i^{normalized} = \frac{V_i - \bar{V}_i}{\hat{\sigma}_{V_i}}. \quad (1)$$

The distributions of normalized features are shown in Figure 3. These features in the Figure 3 are ordered as per the Table 7 in the Appendix.

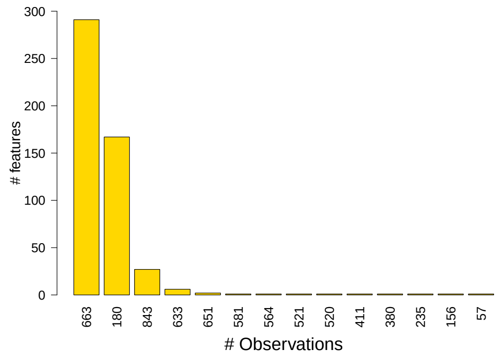

Figure 1: The Frequency of Observations of Different Feature Variables.

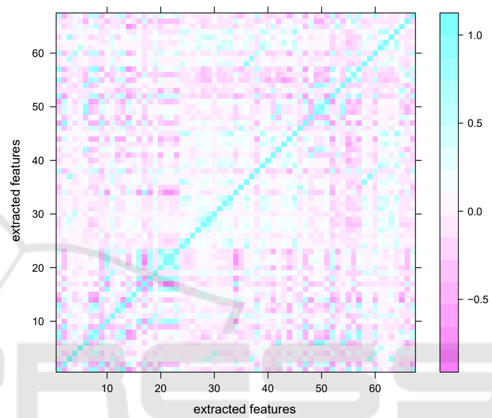

Figure 2: Correlation between Different Features of Process Variables after Filtering.

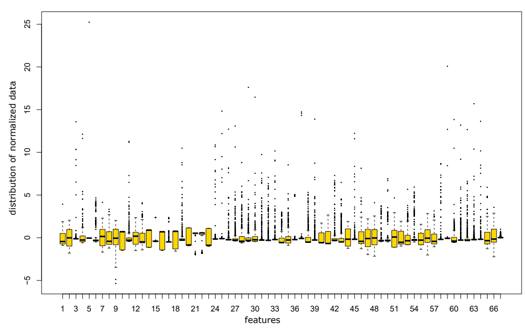

Figure 3: Distributions of Normalized Features.

4 MODEL AND VARIABLE SELECTION

For the selection of the most important features from our data, we applied three different models, which are beta regression with a backward selection procedure, beta boosting approach with regularization and SVM regression with the radial kernel. The details of the methods are provided below.

4.1 Beta Regression Model

This model is proposed by Ferrari and Cribari-Neto (Ferrari and Cribari-Neto, 2004). The beta regression model is used to predict rates and proportions where the prediction variables $y \in (0,1)$. The underlying assumption of the beta regression model is that the response variable is beta distributed. In such cases, the linear models are not useful for two reasons:

- The model parameters are interepreted with respect to the transformed response $\tilde{y} = log(y/(1-y))$ and not with the real response y.
- The heteroskedastic nature of the data.

The beta distribution in terms of μ and ϕ is expressed as follows:

$$f(y;\mu,\phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)(\phi-1)} \quad (2)$$

where, $\mu = p/(p+q)$ and $\phi = p+q$ ($p, q > 0$ real parameters of the distribution, $\Gamma(x)$ Gamma function). The expected value and variance are $E(y) = \mu$ and $Var(y) = \mu(1-\mu)/(1+\phi)$. ϕ is a precision parameter.

In the beta regression model, we suppose to have $y_1, y_2, \ldots, y_n$ random samples where $y_i \sim B(\mu_i, \phi)$, $i = 1, 2, \ldots, n$. The beta regression model is described as follows:

$$g(\mu_i) = x_i^T \beta \quad (3)$$

with a link function $g(x)$, which can be *logit*, *probit* or $log-log$ link. The expected value μ_i is described as follows:

$$\mu_i = g^{-1}(x_i^T \beta) \quad (4)$$

In order to remove the bias of maximum likelihood estimates of parameters an extension by introducing a regression structure on the precision parameter ϕ can be used (Simas et al., 2010):

$$g_1(\mu_i) = x_i^T \beta \quad (5)$$

$$g_2(\phi_i) = z_i^T \gamma \quad (6)$$

The logit, probit and log-log link function are defined as follows:

$$\text{logit: } g(\mu) = log(\frac{\mu}{1-\mu}) \quad (7)$$

$$\text{probit: } g(\mu) = \Phi^{-1}(\mu) \quad (8)$$

The $\Phi(.)$ is a standard normal distribution function

$$\text{log-log: } g(\mu) = -log(-log(\mu)) \quad (9)$$

The improved *beta regression model* allows non linear predictors for $g_1(\mu_i)$ and $g_2(\phi_i)$ and also gives the bias corrected estimate of maximum likelihood. For our data analysis we used the *betareg* R package (Cribari-Neto and Zeileis, 2010).

4.2 Beta Boosting Model

The boosting approach of beta regression model (Schmid et al., 2013) is based on generalized additive models for location, scale and shape (*GAMLSS*) approach (Thomas et al., 2018; Buehlmann and Hothorn, 2007; Rigby and Stasinopoulos, 2005). The beta boosting model uses the *gamboostLSS* boosting algorithm for the variable selection (Hofner et al., 2016). The brief explanation of beta boosting regression is described in Algorithm 1.

Algorithm 1: Betaboosting Algorithm.

B = 10000
Set iterator itr = 1
Set the initial parameter, $\beta = 0$ and $\gamma = 0$ (values for $g_1(\mu)$ and $g_2(\phi)$).
repeat
Keep the γ fixed and select predictor variables by considering the mean model described in Equation 5.
Update β_i coefficient for which the predictor variable, X_i, improves the beta log likelihood estimate.
Keep β fixed and select predictor variables considering precision model described in Equation 6.
Update γ_i coefficient for which the predictor variable, Z_i, improves the beta log likelihood estimate.
itr = itr +1
until itr = B

4.3 SVM Regression

The support vector regression is based on the support vector machine concept (Drucker et al., 1996; Vapnik, 1995). In SVM regression the input data is mapped to the the high dimensional feature space using non-linear mapping, which is described as follows: Let $f : X \rightarrow \mathbb{R}$

$$f(x) = < w, \phi(x) > +b \tag{10}$$

The $\phi(x)$ is a high dimensional feature space. By using a kernel trick, a kernel function can be used to calculate the inner product in feature space, which is described as follows:

$$\hat{f}(x) = \sum_{i=1}^{n} \alpha_i k(x_i, x) + b \tag{11}$$

Where $\alpha = (K + \lambda I)^{-1} y$. In our regression model we use the *radial kernel function*, which is described as follows:

$$k(x, x^{'}) = exp(\frac{||x - x^{'}||}{2\sigma^2}) \tag{12}$$

4.4 Backward Feature Selection for Beta Regression

For the backward feature selection in beta regression model, we followed a bootstrapped approach where we generate 500 bootstrapped datasets, $D = D^1, D^2 \ldots, D^{B=500}$. For each bootstrap data, we model the beta regression model M^b, where $b = 1, 2, \ldots, 500$. We estimate the parameters and compute the weight of each parameter in the model as follows:

$$w_i = \sum_{b=1}^{B} I(p(M_i^b)) <= \alpha) \tag{13}$$

where $i = 1, 2, \ldots n$. The function $p(M_i^b)$ returns the p-value for parameter i from model M^b and α is the defined significance level. We discard the variable which is least weighted and repeat the analysis until the weights of remaining variables are greater then a certain threshold. In our analysis we set this threshold, $thr = 0.9$. The detailed description of feature selection is shown in Algorithm 2.

4.5 Feature Selection in SVM Regression

For the feature selection in SVM regression, we applied the recursive feature elimination (RFE) method to find the most important features which predict the outcome with higher accuracy. We applied the RFE algorithm with resampling, which iteratively rejects the weakest predictor variable. In our analysis, for each iteration we first tune hyperparameters $\sigma = \{.01, .05, .1 \ldots, 1\}$ and box constraint $C = \{0.01, .1, .5, 1, 2, 4, 8, \ldots, 512, 1024\}$ and then we perform the RFE on the best model by tuning hyperparameters.

Algorithm 2: Backward Selection using Beta Regression.

Set $B = 500$
Set thr = 0.9
Set $\alpha = 0.05$
Set D is the training dataset of $n \times m$ dimension.
repeat
  $D^1, D^2, \ldots, D^{500}$ are 500 bootstrapped datasets.
  $M^1, M^2 \ldots M^{500}$ beta regression models for the bootstrapped datasets.
  Let $p(M_i^b)$ be a function that returns the $p-value$ of the parameter i in the model M^b.
  A is a vector of size m.
  for t = 1 to m **do**
    $k = \sum_{b=1}^{B} I(p(M_i^b) <= \alpha)$
    $A[t] = k/B$
  end for
  Let $i = r(A)$ ▷ returns the lowest rank index of parameter
  if $A[i] < thr$ **then**
  $D = D_{\setminus i}$ ▷ Discard the variable i in data D
  else
  $i = NULL$
until $i = NULL$
return A ▷ index of selected parameters.

5 RESULTS

5.1 Betaregression Model

By applying Algorithm 2, we obtained eight important features which are significant consistently in the backward selection. In Figure 4 we show the boxplots of R^2 measures as we discard the weakest variables in each step. The x-axis displays the number of variables in different models, and the y-axis shows the distribution of R^2 values from different models built on the bootstrapped datasets. The R^2 results show gradually decrease as we discard a variable in each step. The average R-square is $\bar{R}^2 = 0.562$ with finally eight variables. Figure 5 visualizes the validation errors and test errors. The validation and test errors decrease as we discard the weakest variable. The error is lowest when there are only eight variables in the model.

5.2 Betaboosting Model

For the betaboosting model we analyze the same training data for different step lengths $S = \{.001, .01, .05, , 0.1, 0.2, 0.5, 1\}$ with 10,000 iterations. Table 1 shows the results for each step length. As performance measures R^2 and the *RMSE* are calculated. The beta boosting model detects

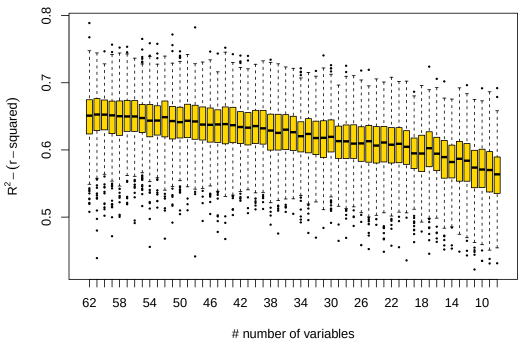

Figure 4: Distribution of R^2 Measures for Different Beta Regression Models using Bootstrapped Samples.

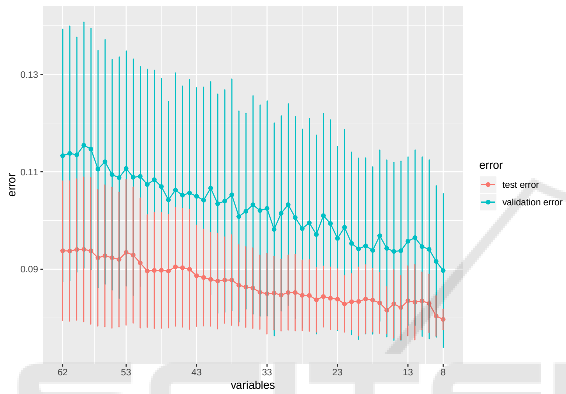

Figure 5: RMSE of Different Beta regression Models using Bootstrapped Samples.

different number of features for μ and ϕ which are shown in Table 1. The best model have higher $R^2 = 0.616$ measure compare to the beta regression model.

5.3 SVM Regression Model

In the first step of the analysis, we train our model with SVM regression using all the feature variables. This model provides improved accuracy in terms of R^2 measure and RMSE. The higher accuracy indicates a non-linear relationship between the scrap rate and the input feature variables. We further applied the RFE algorithm and tuned the model with ten-fold cross-validation with various combinations of hyper-parameters. The results of the best performing models are shown in Table 2. The SVM model recognizes 42 feature variables as important variables.

5.4 Comparison of Models

The three different models identify a different number of features as important predictors for the *scrap rate*. The simplest beta regression model predicts 8 feature

Table 1: R^2 measure and number of significant features for μ and ϕ in Betaboosting regression models using different step lengths.

Step length	R^2	# of features for μ	# of features for ϕ	RMSE
0.001	0.51	11	7	0.103
0.01	0.58	27	17	0.0885
0.05	0.605	40	36	0.0857
0.1	0.616	45	54	0.0846
0.2	0.607	42	36	0.0856
0.5	0.613	27	39	0.0847
1.0	0.597	14	19	0.090

Table 2: The results of best performing SVM models with different combinations of hyper-parameters.

SVM Models	Sigma	C	Rsquared	RMSE
1	0.050	4	0.675	0.0758
2	0.010	32	0.671	0.0757
3	0.010	16	0.667	0.0774
4	0.010	64	0.657	0.0775
5	0.100	4	0.655	0.0778
6	0.050	16	0.654	0.0777
7	0.005	64	0.653	0.0790
8	0.050	2	0.649	0.0809
9	0.100	2	0.646	0.0798
10	0.005	32	0.646	0.0808

variables; the beta boosting model predicts 45 feature variables and SVM regression detects 42 important feature variables. 6 feature variables are common in all three models, 7 feature variables are common in the beta regression model, and the SVM model. The 31 feature variables are common in the beta boosting model and SVM regression. The common 6 variables, which are present all three models, are shown in Table 3. Apart from the common features, different features are ranked high by the beta boosting model and SVM model. These top features are shown in Table 5 for each approach. The reason for different feature weights depends on the underlying assumptions of different models; therefore different features are weighted high by different models.

We further test these three models on the testing data, which consists of 160 testing samples, the RMSE of the different models are shown in Table 4. The predicted scrap rates for the testing data are shown in Figure 6 for each applied model. By analyzing these three models, we found that up to 65% of the variance in scrap can be described by at least ~ 42 features of process variables as shown by the SVM model. However, a large percentage of variance in the output is described by 8 feature variables resulting in the beta regression model.

Table 3: Common Process Variables and their most Important Features.

Process variables	Features
Screw volume (end of holding-pressure phase)	$\bar{X}, S(X), M_1(x)$
Torque peak value	$max(x)$
Temperature zone 7	$max(x)$
Temperature zone 11	$max(x)$

Table 4: RMSE on Testing Data for Different Models.

Models	Beta Regression	Beta Boosting	SVM Regression
RMSE	0.0792	0.0793	0.0764

The nonlinear models do not add more value to the results, which leads to the conclusion that a small number of variables are linearly related to the scrap rate. Also, the processed data do not explain the remaining variance in the data. The RMSE of the testing data suggests that the SVM model leads to the lowest prediction error. However, the testing error of the Beta regression model is only slightly higher than the SVM model. The testing data show a higher variance for some of the process segments, and this is common in all models for different scrap rates. The high variance in the prediction when scrap is higher can be due to some technical faults, which are unnoticed or not recorded. Similarly, an unexplained variance in the scrap rate prediction by predictive models also have a dependency on other features such as material type, product shape, and other external factors.

5.5 Comparison with Related Studies

We compare our approach with two other approaches shown in Table 6. The first difference is that the data in these two approaches is generated from an experimental design by controlling fewer (≤ 6) process variables on a single machine. The data is generated by controlling fewer process variables of different categories of products quality. These methods utilize machine learning approaches for the classification of product quality using SVM and deep-learning approaches. In our analysis, we predict the scrap rate as an output of the whole production process by using statistical features of process variables, which is a high dimensional data of ~ 65 feature variables. Also, we do not differentiate between different types of scraps. The results in both the studies predict scrap class with higher accuracy, but they are not directly comparable with our results due to the nature of the data and the output. However, our models predict some common features which show importance in previous studies such as cycle time, screw volume, torque, different temperature zones and pressure parameters (Singh and Verma, 2017).

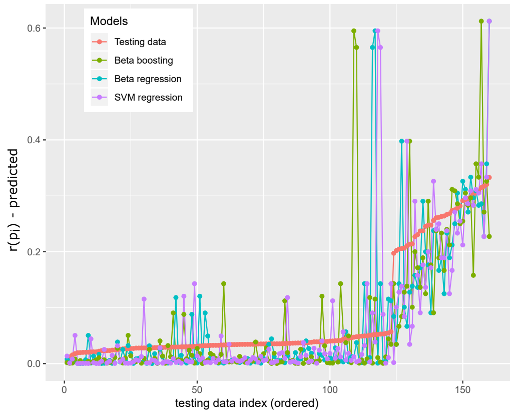

Figure 6: Predicted r_{p_i} by different models on testing data.

Table 5: Top Process Variables ranked by Different Models.

Process variables	Features
Beta Regression	
Screw volume (end of holding-pressure phase)	$M_1(X), S(X), \bar{X}$
Discharge end 1	$max(X), M_1(X)$
Torque peak value	$max(X)$
Temperature zone 7	$max(X)$
Temperature zone 11	$max(X)$
Beta boosting	
Screw volume (end of holding-pressure phase)	$S(X)$
Torque peak value	$max(X), min(X)$
Temperature zone 6	$max(X)$
Shot volume	$max(X)$
discharge end 1	$S(X)$
Integral monitoring micrograph	$max(X)$
Temperature zone 7	$M_1(X)$
Hydraulic pressure at switch point	$M_1(X)$
Idle time before cycle start	$M_1(X)$
SVM Regression	
Screw volume (end of holding-pressure phase)	$S(X)$
Shot volume	$max(X)$
Temperature zone 7	$min(X)$
Temperature zone 8	$max(X)$
Torque peak value	$max(X), min(X)$
Actual value pressure pump	$M_1(X)$
Cycle time	$M_1(X), max(X)$
Actual value injection time	$min(X)$
Idle time before cycle start	$M_1(X)$
Temperature zone 6	$max(X)$

6 CONCLUSIONS

In this paper, we analyzed production data from an injection molding process, which contains ~ 70 process variables from different machines. We start our analysis by preprocessing, cleaning and filtering of relevant information from the raw data. We extracted important statistical features of different process variables. After filtering using Algorithm 3, we selected 66 of them. These 66 features are used for scrap rate prediction using linear and non-linear models. We first applied the beta regression model with a backward feature selection method, which provides a significant estimate of the scrap rate. We extend it for the non-linear models to explore the non-linear relationship between the scrap rate and the feature variables. The non-linear models provide a slight improvement in prediction, but due to a large number of selected features, the models become more complex.

In this analysis, we try to understand more general, which feature variables affect the quality of the production process. A simple beta regression model provides a good prediction for the scrap rate with fewer feature variables than the non-linear models. However, the unexplained variance of the response variable also depends on many other aspects of the product such as the material type, size, volume and many other product specific features, which are not the part of the analysis. Additionally, there are many machine-specific features which have been discarded in our prepossessing due to not having enough samples and variance. The product and machine specific features which are absent in the model can be the reason for the unexplained variance of the predictive models. The features of the process variables, identified by different methods, are general features which affect the product quality of different product types. Particular attention should be paid to tuning of these process parameters for a better production quality. However, the product quality also depends on material types, volume and size, and other product and machine specific parameters. Therefore, in our future work, for the product-specific quality control, we will look into the data in more details by exploring more product and machine specific details which affect the production output. We want to extend our data modeling for the different type of product specific quality measures.

COPYRIGHT FORM

The Authors have no competing interest.

Table 6: Comparison of three different approaches for the prediction of product quality.

Methods proposed by	Objective and Methodology	Data generation	Input variables and Output
(Ribeiro, 2005)	Product quality prediction using Support vector machine based approaches by tuning different hyperparameters for error classification	Data is generated by an experimental set up on Demag injection molding machine with Hostacom DM2 T06 polymer and with mold DN502	Input variables: cycle time, dosage time, injection time, cushion, peak melt temperature and, ram velocity Output: Product quality of different categories which are Streak, Strains, Burn marks, Edges, Unfilled parts and Warped parts
(Mao et al., 2018)	Feature learning and process monitoring using Deep learning approach of Convolution-deconvolution auto encoder	Data is generated by an experimental set up on on a *JSW J110ADC-180H* electric injection molding machine by tuning different process conditions to generate different batches of good and faulty quality.	The data is in the form of 4D input Tensor, $\chi(B \times V \times T \times C)$, where, V is the number of variables, B is the batch size of a product quality class, C is the number of feature channels, and T is the set of time instances where the values of different process variables are measured. V={screw displacement, injection pressure, cavity pressure } Output: Different conditions of product quality
Our approach	Identification of key process variables which affect the product quality (scrap rate) using Beta regression, beta boosting and SVM regression methods	Observational data from real time production of different products of shape, size and material type produced by 33 different machines in the company.	65 Statistical features extracted from different process variables Output: Scrap rate which is the proportion of total scraps and the total output produced.

ACKNOWLEDGEMENTS

This paper was funded through the project ADAPT by the Government of Upper Austria in their programme "Innovative Upper Austria 2020".

REFERENCES

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

Buehlmann, P. and Hothorn, T. (2007). Boosting algorithms: Regularization, prediction and model fitting. *Statist. Sci.*, 22(4):477–505.

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

Chen, W.-C., Nguyen, M.-H., Chiu, W.-H., Chen, T.-N., and Tai, P.-H. (2016). 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.

Cribari-Neto, F. and Zeileis, A. (2010). Beta regression in r. *Journal of Statistical Software, Articles*, 34(2):1–24.

Dang, X.-P. (2014). General frameworks for optimization of plastic injection molding process parameters. *Simulation Modelling Practice and Theory*, 41:15 – 27.

Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A., and Vapnik, V. (1996). Support vector regression machines. In *Proceedings of the 9th International Conference on Neural Information Processing Systems*, NIPS'96, pages 155–161, Cambridge, MA, USA. MIT Press.

Fernandes, C., Pontes, A. J., Viana, J. C., and Gaspar-Cunha, A. (2018). Modeling and optimization of the injection-molding process: A review. *Advances in Polymer Technology*, 37(2):429–449.

Ferrari, S. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7):799–815.

Hofner, B., Mayr, A., and Schmid, M. (2016). gamboostlss: An r package for model building and variable selection in the gamlss framework. *Journal of Statistical Software, Articles*, 74(1):1–31.

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

Kashyap, S. and Datta, D. (2015). Process parameter optimization of plastic injection molding: a review. *International Journal of Plastics Technology*, 19(1):1–18.

Mao, T., Zhang, Y., Ruan, Y., Gao, H., Zhou, H., and Li, D. (2018). Feature learning and process monitoring of injection molding using convolution-deconvolution auto encoders. *Computers & Chemical Engineering*, 118:77 – 90.

Oliaei, E., Heidari, B. S., Davachi, S. M., Bahrami, M., Davoodi, S., Hejazi, I., and Seyfi, J. (2016). 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.

Packianather, M., Griffiths, C., and Kadir, W. (2015). Micro injection moulding process parameter tuning. *Procedia CIRP*, 33:400 – 405. 9th CIRP Conference on Intelligent Computation in Manufacturing Engineering - CIRP ICME '14.

Ribeiro, B. (2005). 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.

Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 54(3):507–554.

Schmid, M., Wickler, F., Maloney, K. O., Mitchell, R., Fenske, N., and Mayr, A. (2013). Boosted beta regression. *PLOS ONE*, 8(4):1–15.

Simas, A. B., Barreto-Souza, W., and Rocha, A. V. (2010). Improved estimators for a general class of beta regression models. *Computational Statistics & Data Analysis*, 54(2):348 – 366.

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

Taguchi, G., Elsayed, E., and Hsiang, T. (1989). *Quality engineering in production systems*. McGraw-Hill series in industrial engineering and management science. McGraw-Hill.

Taguchi, G., Konishi, S., Konishi, S., and Institute, A. S. (1987). *Orthogonal Arrays and Linear Graphs: Tools for Quality Engineering*. Taguchi methods. American Supplier Institute.

Thomas, J., Mayr, A., Bischl, B., Schmid, M., Smith, A., and Hofner, B. (2018). Gradient boosting for distributional regression: faster tuning and improved variable selection via noncyclical updates. *Statistics and Computing*, 28(3):673–687.

Unal, R. and Dean, E. B. (1991). Taguchi approach to design optimization for quality and cost : An overview.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag, Berlin, Heidelberg.

Yin, F., Mao, H., and Hua, L. (2011a). A hybrid of back propagation neural network and genetic algorithm for optimization of injection molding process parameters. *Materials & Design*, 32(6):3457 – 3464.

Yin, F., Mao, H., Hua, L., Guo, W., and Shu, M. (2011b). Back propagation neural network modeling for warpage prediction and optimization of plastic products during injection molding. *Materials & Design*, 32(4):1844 – 1850.

Zafošnik, B., Božič, U., and Florjanič, B. (2015). 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.

APPENDIX

Statistical Features

$$\bar{X} = \frac{x_{t_1} + x_{t_2} + \cdots + x_{t_n}}{n} \quad (14)$$

$$S(X) = \sqrt{\frac{\sum_{i=1}^{n}(x_{t_i} - \bar{X})^2}{n-1}} \quad (15)$$

$$min(X) = min(x_{t_1}, x_{t_2}, \ldots x_{t_n}) \quad (16)$$

$$max(X) = max(x_{t_1}, x_{t_2}, \ldots x_{t_n}) \quad (17)$$

$$M_1(X) = \frac{max(X) - min(X)}{2} \quad (18)$$

$$M_2(X) = \frac{max(X)}{\bar{X}} \quad (19)$$

Data Filtering Algorithm

Algorithm 3: Filtering out the features which have correlation, $\rho \geq .95$.

Initialize:
R is a $p \times p$ sample correlation matrix
$V = \{V_1, V_2, \ldots, V_p\}$ is a ordered set of features based on the number of other features with which it has correlation, $\rho \geq .95$, which is calculated as follows.

$$a(V_i) = \sum_{k=1}^{(p-1)} I(\rho(i, k_{\setminus i}) \geq 0.95)$$

$V = \{V_1, V_2, \ldots, V_p\}$, is ordered by $a(V_1) > a(V_2) > \cdots > a(V_p)$
L is an empty vector.
$cnt = 1$
repeat
    $L[cnt] = V_1$
    $V' \subset V$ is a set of features with which $R[V_1, V'] \geq= .95$
    Update V: $V = V \setminus V'$
    cnt = cnt + 1
until V is empty
return L

Features After Data Filtering

In Table 7 the subset of 66 variables after filtering using Algorithm 3 can be found. The extracted features f1, f2, …, f6 refer to Eq. 14, 15, …, 19. The respective process variable comes after the "_" sign.

Table 7: Selected feature variables after filtering.

S. no.	Features
1	f4_schussvolumen_1
2	f1_tempzone_3_istwert
3	f6_drehzahl_spitzenwert_host_1
4	f5_schneckenvolumen_ende_nachdruckcpnschneckenposition_ende_nachdruck_1
5	f5_fliesszahl_1
6	f4_hydr_druck_beim_umschalten_1
7	f4_entlastung_ende_1
8	f4_drehmoment_spitzenwert_laufender_zyklus_host_1
9	f3_tempzone_3_istwert
10	f1_tempzone_12_istwert
11	f5_tempzone_5_istwert
12	f4_zykluszeit_sollwert
13	f4_tempzone_8_istwert
14	f4_tempzone_7_istwert
15	f4_tempzone_6_istwert
16	f4_tempzone_11_istwert
17	f4_integral_ueberwachung_2_micrograph
18	f3_tempzone_11_istwert
19	f2_tempzone_5_istwert
20	f1_tempzone_9_istwert
21	f1_tempzone_14_istwert
22	f1_tempzone_13_istwert
23	f1_tempzone_10_istwert
24	f6_tempzone_9_istwert
25	f6_tempzone_14_istwert
26	f6_staudruck_spitzenwert_1
27	f6_spez_nachdruck_spitzenwertpnshydr_nachdruck_spitzenwert_1
28	f6_spez_einspritzdruck_spitzenwertpvshydr_einspritzdruck_spitzenwert_1
29	f6_spez_druck_beim_umschaltenphuhydr_druck_beim_umschalten_1
30	f6_schneckenvolumen_ende_nachdruckcpnschneckenposition_ende_nachdruck_1
31	f6_schliesskraft_spitzenwert
32	f6_entlastung_ende_1
33	f6_dosiervolumenswc1dosierhub_1
34	f6_aktuelles_umschaltvolumenc3uaktuelle_umschaltposition_1
35	f5_zykluszeit_vollautomatik
36	f5_tempzone_9_istwert
37	f5_tempzone_7_istwert
38	f5_stillstandszeit_vor_zyklusstart
39	f5_schliesskraft_spitzenwert
40	f5_ruesten
41	f5_mengenistwert_pumpe
42	f5_hydr_druck_beim_umschalten_1
43	f5_entlastung_ende_1
44	f5_druckistwert_pumpe
45	f5_dosiervolumenswc1dosierhub_1
46	f4_zykluszeit_vollautomatik
47	f4_tempzone_2_istwert
48	f4_tempzone_1_istwert
49	f4_staudruck_spitzenwert_1
50	f4_nachdruck_spitzenwert_1
51	f4_integral_ueberwachung_1_micrograph
52	f3_spritzzeit_istwert_1
53	f3_schneckenvolumen_ende_nachdruckcpnschneckenposition_ende_nachdruck_1
54	f3_hydr_druck_beim_umschalten_1
55	f3_entlastung_ende_1
56	f3_drehzahl_spitzenwert_host_1
57	f3_drehmoment_spitzenwert_laufender_zyklus_host_1
58	f2_tempzone_9_istwert
59	f2_tempzone_7_istwert
60	f2_schneckenvolumen_ende_nachdruckcpnschneckenposition_ende_nachdruck_1
61	f2_schliesskraft_spitzenwert
62	f2_entlastung_ende_1
63	f2_drehzahl_spitzenwert_host_1
64	f2_dosiervolumenswc1dosierhub_1
65	f1_schneckenvolumen_ende_nachdruckcpnschneckenposition_ende_nachdruck_1
66	f1_drehzahl_spitzenwert_host_1