Time Series Classification for Scrap Rate Prediction in Transfer Molding

Anna Mándli[1,3] Róbert Pálovics[2] Mátyás Susits[3] András A. Benczúr[2]

[1]Robert Bosch Kft, Engineering Center Budapest
[2]Institute for Computer Science and Control, Hungarian Academy of Sciences
[3]Eötvös Loránd University, Budapest

anna.mandli@hu.bosch.com, {rpalovics, susitsm, benczur}@info.ilab.sztaki.hu

ABSTRACT

In this paper, we present and evaluate methods for predicting critical increase in manufacturing scrap rate of automotive electronic products. Along with information on processes such as maintenance cycles, we analyze the sensor time series of the so-called transfer molding process, in which the electronic product is packaged into plastic for protection. Production data are organized in a two level hierarchy of the individual parts and of the sequence of parts. Since the main goal is to predict and warn about the future state of the process, we designed a training and prediction framework over certain production cycles. By using sensor and other information, we adapt known time series classification methods to predict increase in scrap rate in the near future. By using three months of manufacturing time series, we evaluate both feature based and dynamic time warping based methods that are capable of fusing a large number of production time series. As a main conclusion, we may warn the operators of increase in failures with an AUC above 0.7 by combining multiple approaches in our final classifier ensemble.

1. INTRODUCTION

Smart factories are considered the next industrial revolution [21] with the main promise that monitoring the sensor data of manufacturing processes real time, we may predict and mitigate failures in the production process [11]. The difficulty of predicting industrial processes lie in the complex structure of the available data. Sensor information from manufacturing is richer compared to typical time series data, since information is organized hierarchically into individual products, leadframes, charges, work shifts, cleaning cycles, etc.

In this paper, we address transfer molding, in which electronic products are packaged in plastic to protect from external damages. During the process, multiple pressure and temperature sensor time series are recorded for each individual product. Finally all these data are arranged in series of the entire product lifetime, segmented by different units of production driven by cleaning cycles, possibly several times interrupted for maintenance and calibration of the molding machine.

Our main target is to reduce the number of failed products during the molding process. At first glance, the solution could be modeling individual product failure based on its time series. While such models may serve for root cause analysis, their output is meaningless, since we will already know if the product failed just minutes after molding. Instead, we focus on predicting the scrap rate in longer segments of the manufacturing process.

In order to warn the operators in case of anticipated production problems, our goal is to predict the increase in scrap rate at certain process cycle starting points. More precisely, whenever the machine goes through cleaning and other possible maintenance and calibration steps, we apply our models to the data of the the next *charge*, a block of products detailed in Section 1.1. We train based on the data of these first charges immediately after cleaning. The classification target is whether the scrap rate is increased above certain threshold in the following charges. Note that evaluation metrics based on the scrap rate itself e.g. squared error are less meaningful, as the scrap rate normally changes little compared to past charges.

We predict increase in scrap rates by a variety of similarity and feature based time series methods. The first and most important step is feature engineering of the set of time series: in the few minutes it takes to produce an individual product, 500 sensor measurements are taken, which are then organized hierarchically in time based on process cycles. We take 50 characteristic time points for each individual product and form 50 time series, where each point corresponds to a product. We compute statistics such as mean, minimum, maximum, variance and derivatives, as well as dynamic time warping (DTW) distance [10] from past production cycles. We use nearest neighbor, SVM [20] and gradient boosted trees (GBT) [6] for prediction.

The observed performance of the methods is summarized as follows. Models based on features of past scrap rates perform very well, however these methods conclude the fact that if the recent scrap rate is high and we observe a lower rate in the present, then it is likely to increase. Since warnings in periods of high scrap rates give no new information to the operators, we consider scrap rate features a baseline. We indicate these in our measurements, but do not otherwise use or combine. The main models are based on support vector regression over time series statistics and 2-nearest-neighbors over DTW. These methods perform similarly, and usually better than the scrap rate baseline. The strongest method is obtained by taking a weighted linear combination of the two model scores. The combination outperformed the baseline in all our measurements. Surprisingly, GBT shows weak

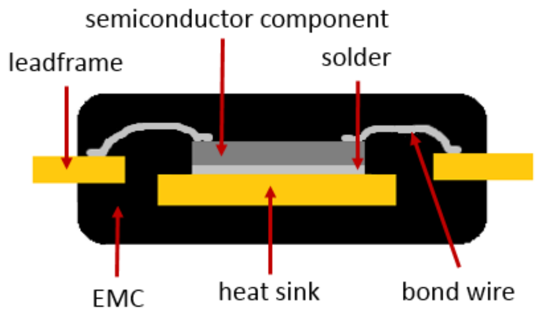

Figure 1: Cross-section of a molded part.

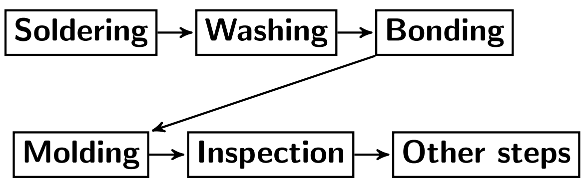

Figure 2: Main steps the assembly process.

performance on all our feature sets.

The paper is organized as follows. In Section 1.1, we detail the transfer molding process. Section 2 describes the available data, Section 3 the framework of our experiments, Section 4 the feature engineering procedure, Section 5 the time series classification methods and Section 6 summarizes the results of our experiments. Finally, related results are summarized in Section 7.

1.1 Transfer Molding

Transfer molding is one among a number of technologies that protect automotive electronic products from external environment (heat, moisture, pressure, vibration). The semiconductor components are packaged using epoxy molding compound (EMC). The schematic cross-section of the automotive electronic part of our consideration is shown in Fig. 1.

Main steps of the assembly process are illustrated in Fig. 2. First, semiconductor components are *soldered* on the copper substrates (heat sinks) with reflow soldering. In the next cleaning step, the products are *washed* with solvents in order to remove flux residues. In the next step, the components are connected by wire *bonding*. From these first three steps, we only consider washing as it has measurable effect on product quality. Data on soldering and bonding will be ignored for now.

In the main *transfer molding* step, the product is packaged using transfer molding by an automold machine. The machine molds several products at once, and the production is organized into units with no change in the production parameters. In order to understand the structure of the measurements, next we describe the production as summarized in Fig. 3.

The machine has *four presses* with *two cavities* in each press. A *shot* is the smallest unit of production involving one press. Data from the presses differ even if they operate simultaneously. In one shot, one *leadframe* per cavity is produced, each of which has three products fixed on them.

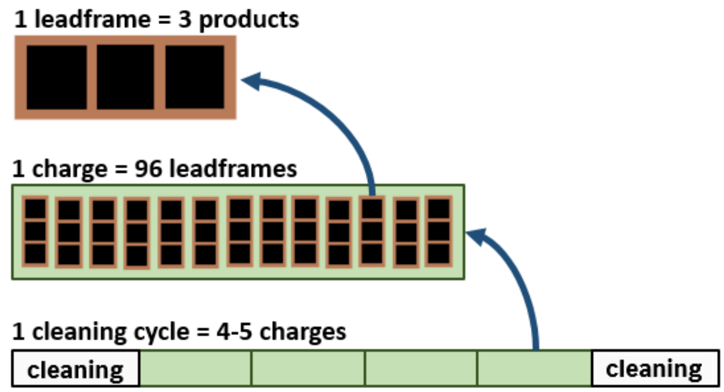

Figure 3: Organization of the products on leadframes, charges, cleaning cycles.

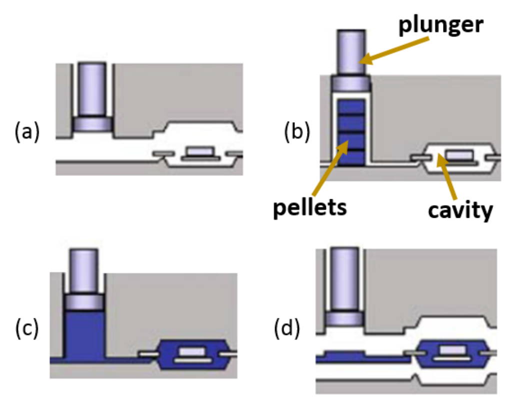

Figure 4: Transfer molding process steps. The image shows a single cavity of one press.

Hence in one shot, two leadframes of altogether six products are molded, and these have the exact same data recorded.

Leadframes are organized into *charges* of continuous operation when no production parameters are changed. One charge contains 96 leadframes with 288 products altogether. Approximately 10 charges are produced daily in three shifts, but the line is used for producing other products as well, hence our data is not continuous in time.

The process steps of transfer molding are summarized in Fig. 4. (*a*) The leadframe (containing the heat sinks and the components) is placed in the preheated mold; (*b*) the mold is closed, and the EMC pellets are fed into the tool; (*c*) the plunger starts to move, and the cavity is filled with melted EMC. After the plunger stops, the EMC is cured under pressure; (*d*) after curing, the mold is opened.

The machine is cleaned regularly after around 4–5 charges by making a shot with conditioner pellets. Before and after cleaning, the machine is idle for several hours, which gives a natural segmentation of the process. Also, changes in scrap rate are expected after the idle periods, either simply because of the possible contamination that remains after cleaning, or of other maintenance operations in the idle period that may affect quality.

After molding, the products are *inspected* for failures. Inspection may be by microscopic pictures of cross sections,

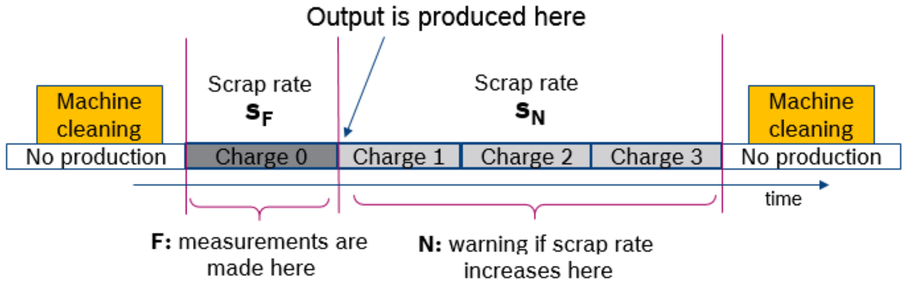

Figure 5: Timeline of a cleaning cycle: between two machine cleanings, four charges were produced (Charge 0-3). The first charge F after machine cleaning is observed. The scrap rate of the charge F is denoted by s_F. The prediction targets the relative increase of the scrap rate in the next N charges until cleaning. The average scrap rate of N charges is denoted by s_N.

or, as in our experiment, by a non-destructive way using scanning acoustic microscope (SAM). At this station, the different failures on the parts are counted and recorded (e.g. upper/bottom cracks, delamination, and voids that we describe later). Further steps of the manufacturing (PMC, milling, trim and form, electrical tests) are not considered in our research.

Molding parameters. Transfer molding is done on high temperature, and the molding time is few minutes. The other important parameter to set is the pressure on the plunger. While the plunger is moving in the first part of step (c) in Fig. 4, the plunger position is controlled. During the curing, both duration and plunger pressure is preset to a constant. The typical shape of the time series recorded when molding a leadframe is shown in Fig. 6.

Failure types. Typical failure types of transfer molding are voids (air inside the EMC), delamination (air between substrate and EMC, or components and EMC), and cracks. The root cause may lie in the design of the part (e.g. molding material cannot fill fully the cavity because of the geometry), or in inappropriate parameters of the process (e.g. too fast or too slow curing, wrong temperature or pressure settings).

In our research, we will predict the rate of *delamination*. Simulation of delamination based on physical models is well studied [19, 23]. These simulations are made in the product development phase in order to find the manufacturing parameters which minimize the chance for delamination. Our model is developed to be used in production, the analysis is data driven and we do not build upon the physical models of transfer molding.

2. DATA

For our research, full manufacturing data of the assembly line was collected for three months. In a discontinuous operation involving maintenance and different products manufactured, approximately 30,000 leadframes of the same product were produced, which we analyze in this paper. In this section, we describe the structure and the variables of the data, which are summarized in Table 1.

Process cycles are defined as periods to predict potential increase in scrap rate, as shown in Fig. 5. A process cycle is a unit of observation consisting of the charges between two machine cleaning steps. We expect changes in behavior after these idle periods of the machine. To gather information on the current state of the machine, we measure all parameters

Process step	Data
Washing	time elapsed between washing and molding bath quality
Molding	timestamp transfer graphs shot count press ID tool position tool ID temperature measurements mold materaial and conditioner lot number
Inspecion	number of delaminations

Table 1: Parts of the data and the process steps where they were recorded.

of the leadframes F in the first charge after cleaning. Our aim is to design a system that warns if scrap rate increases in the subsequent charges N. Since we consider each press separately, we have 164 process cycles, each of which will be used for giving a prediction on F and evaluating on N.

Washing related features are time elapsed between washing and molding, and *cleaning bath quality* that is expressed as the number of hours elapsed since the last washing bath change.

Transfer graphs. The time series of measurements recorded during molding of a leadframe are often called *transfer graphs*. The studied machine records three transfer graphs of 500 measurements for each shot. These are the position of the plunger, pressure (calculated from force measurements on the plunger) and vacuum, see Fig. 6.

An important phase of molding is the time period when the plunger moves, and the EMC is flowing. We will refer to the pressure measurements $1-50$ which were made during this time as *filling pressure*.

Other transfer molding process data consists of timestamp of molding, raw material (mold compound and conditioner) lot number, temperature measurements, and shot count since the last cleaning. The position where the leadframe was molded inside the machine is described by press ID and tool position, since the machine has four presses with two cavities (tool positions) in each. Tools inside the cavities are sometimes changed, this can be tracked by the variable tool ID.

Result of the visual inspection is the number of de-

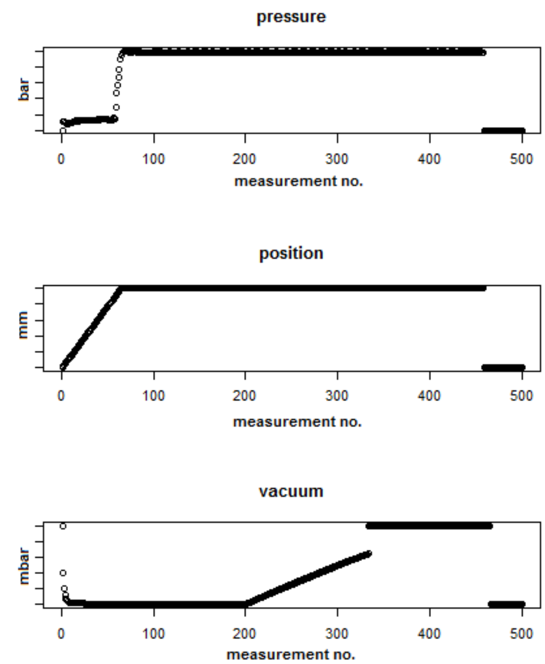

Figure 6: Transfer graphs: three time series of 500 measurement points recorded during the transfer molding process. Scales on the y axes are hidden for confidentiality reasons.

tected delaminations for all parts. From this, we will only use the rate of delaminated leadframes in certain charges or group of charges, and these rates are referred to simply as scrap rates later on.

3. TRAINING AND EVALUATION

Since manufacturing data is identical for the products on the same leadframe, we consider them as one record. A leadframe is defined to be scrap if it has at least one delamination. For a set of leadframes L, the *scrap rate* s_L is the fraction of leadframes with at least one delamination.

The average scrap rates of the data can be seen in Fig. 7. Based on these the data can be roughly separated into three intervals over time. The first interval has the lowest, while the last has the highest average scrap rate.

In order to qualify the increase in scrap rate, we define both an additive and a multiplicative constant. For very low scrap rates, a relative increase may have no practical relevance and the case may be similar for small additive changes in case of high scrap rates. With the notion of first and next charges F and N, the *target label* of the charge is 1 if $s_N - 1.05 \cdot s_F > 0.02$. We selected the parameters 1.05 and $t = 0.02$ such that the positive target has business relevance and occurs both for the good and bad production phases, see Fig. 7.

The purpose of our testing procedure is to simulate a warning system based on different classifiers. Recall that a process cycle consists of charges of the same press between two cleaning steps. For each process cycle, we train on data of earlier ones. The model is applied to data of the first charge F and evaluated on N subsequent charges before the next cleaning, as seen in Fig. 5. We evaluate by computing

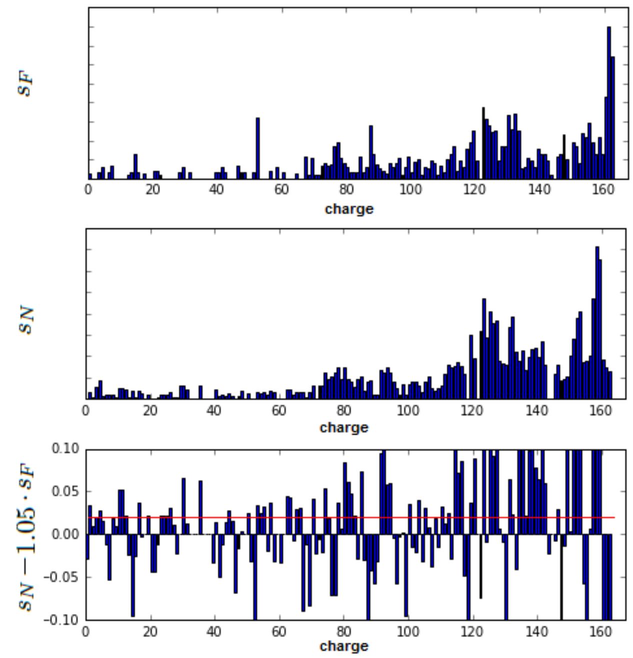

Figure 7: *Top:* Scrap rate of the charges following the machine cleaning (s_F). *Middle:* Scrap rates of the charges where the increase of the scrap rate should be predicted (s_N). *Bottom:* The function $s_N - 1.05 \cdot s_F$, which is used to generate the binary target with a threshold, in our case 0.02 (horizontal line). Charges are in production order, the scale for scrap rates is hidden for confidentiality reasons.

the raw prediction score of the models, which we consider as rank for computing the AUC of the procedure. By the mathematical properties of AUC [5], we obtain the probability that a random process cycle with increased scrap rate is prioritized higher for warning than another random one with no increase. Tthis way the warning threshold may be left as a free parameter for the machine operator.

4. FEATURE ENGINEERING

Recall the hierarchical arrangement of products into shots, charges and cleaning cycles as summarized in Figs. 3 and 5. The organization of the production makes our feature engineering task more cumbersome than in typical time series analysis tasks. Recall that a process cycle consists of charges between two machine cleaning steps, and in one charge of identical machine configuration, 48 shots of 96 leadframes with altogether 288 products are molded. Our most important task is to define charge level features based on the individual shot time series and other process information that we list next.

Past scrap rate time series give a highly accurate prediction by themselves. However, this type of prediction lacks novelty to the operator. We generate moving averages to define baseline classifiers, that will only be used in comparison with the predictions of other methods.

Transfer graph statistics. First, we chose features which can efficiently describe the individual transfer graphs. For example, the position graph consists of straight seg-

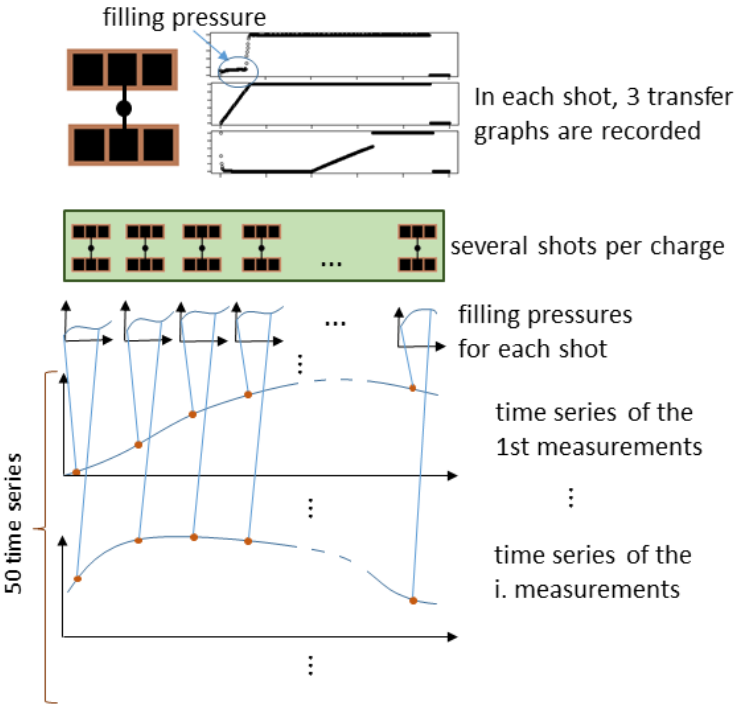

Figure 8: The structure of the set of sensor time series. In a single charge (see Fig. 3), we form 50 time series of the same characteristic points of the individual shot time series.

ments, and thus we can describe it with the slope of the first segment and with the maximum. Charge level features were obtained from these by taking the mean (minimum, maximum) for the leadframes of the given charge.

Filling pressure by product series characterize how the most important measurement, the filling pressure changes as the series of products are molded. The first 50 measurements of each product, as seen in the top of Fig. 6, are arranged in 50 time series that characterize a charge, as seen in Fig. 8. We consider the matrix p_{ij} for measurement i on shot j. Instead of considering the time series $\{p_{1j}, p_{2j}, \ldots\}$ of each shot j, we produce the time series $\{p_{i1}, p_{i2}, \ldots\}$ from several shots where i is a characteristic point of the transfer graph. In our experiments we use 50 different transfer graph points.

As an illustration of the product series, in Fig. 9 we see that the filling pressure values are lower after machine cleaning, and they typically increase later on.

Feature importance was calculated by predicting individual product failure using gradient boosted tree classifiers [6]. We keep the relevant features, which consist of features of the filling pressure and the vacuum graph, the shot count since cleaning, and information on washing.

Removed features are in trivial connection with time and hence overfit to the three periods of production with different scrap rates. The list of these variables is the following. Vacuum maximum is equal to the outside the air pressure and identify the time of production very well. The value of the first few vacuum measurements is also strongly correlated. The slope of the vacuum transfer graph, although

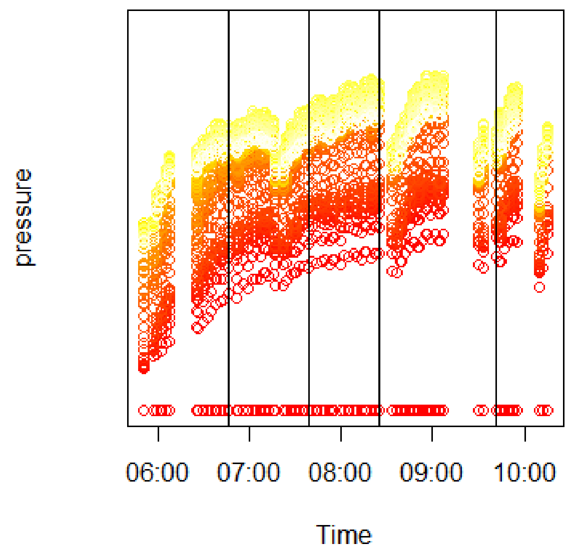

Figure 9: Filling pressure measurements. Red to yellow color scale indicates measurement points 1 through 50. Vertical lines indicate new charges.

relevant, has a value that increases slowly in time as the vacuum chamber quality decreases and maintenance is performed on the chamber only a few times a year. The maximum of the position transfer graph depends on the presses and the pellet size. The pellet size, although relevant, was only changed once in our period of observation. Most of these variables will be used in a study with extended time frame of data collection.

5. MODELS

We use nearest neighbor, gradient boosted tree and support vector machine classifiers to predict increase in the scrap rate on different feature sets. In this section we describe our methods organized by the type of the features used.

First we recall our unusual evaluation procedure from Section 3. We compute AUC [5] by ranking the process cycles. For each cycle, we compute new models based on past data, which give raw prediction scores. In case of classification models we directly use these raw scores for AUC computation.

We considered regression models where we used the next scrap rates s_N (see Fig. 5) for training. To compute ranking based AUC metric, we may either use the *raw predicted scores* by regression or the *transformed scores* by using the formula $raw - 1.05 \cdot s_F$. Note that this is the same as the formula used to obtain the relative target from scrap rates s_N and s_F which is described in Section 3.

Past scrap rate features. We used this feature set to generate baseline models. These models perform quite well, but predictions on scrap rate using scrap rates of the near past are lacking added value. On the other hand, models using purely the process data are more likely to give warnings

Method	AUC
SVR on time series features (transformed)	0.70
SVR on scrap rates (transformed)	0.68
2-NN DTW (transformed)	0.67
2-NN DTW	0.63
GBT on scrap rates	0.61
SVC DTW	0.59
SVR on scrap rates	0.59
SVR on time series features	0.59
GBT on time series features	0.54

Table 2: AUC for the predictions on the label $s_N - 1.05 \cdot s_F > 0.02$. Best methods are in bold, strongest baseline is in italics.

for unexpected scrap rate increase to the operators.

We used *gradient boosted trees* (GBT) and *support vector regression* (SVR) on the features generated from the scrap rates of the near past. In case of GBT, binary labels could be used directly as target. In case of SVR, we use the values of s_N as target and compute AUC by raw or transformed scores.

Time series features. We use SVR and GBT to predict based on time series statistics of the 50 time series as detailed in Section 4, and illustrated in Fig. 8. The target and the method for calculating AUC is the same as above.

Charge DTW similarity. Dynamic Time Warping (DTW) [10] finds an optimal alignment between the two time series by minimizing the sum of squares of the distances of aligned points. We compute DTW over the 50 series defined Section 4. For two charges S and T, let the 50 time series of the 50 characteristic points be S_i and T_i each, for $i = 1..50$. For DTW distance based classification, a simple but accurate classifier is the *nearest neighbor*, as measured in [22]. For the nearest neighbor algorithm, the distance of charges S and T is obtained as the Euclidean distance $d_{TS} = \sqrt{\sum_{i=1}^{50} dist_{DTW}(T_i, S_i)^2}$. We compute nearest neighbor regression of the next scrap rates (s_N) for the k nearest charges as raw score, which we also transform by $raw - 1.05 \cdot s_F$.

Finally, we also use *support vector classification* (SVC) [20] over DTW distances based on the ideas in [2, 4, 7]. We set a random sample r of charges aside and for the rest of the charges we compute the DTW distances from each charge in r. For each charge, hence we obtain $50 \cdot r$ values as features and we apply support vector machines to classify the increased scrap rate target.

6. EXPERIMENTS

In our experiments, predictions were made by using the models described in Section 5. The prediction target is $s_N - 1.05 \cdot s_F > 0.02$. The AUC values are summarized in Table 2 in descending order. Recall that AUC is computed for a ranking based on the output score of different models. We noticed that the transformed scores $raw - 1.05 \cdot s_F$ perform well, since they adjust the scores to denote the scrap rate increase and not actual scrap rates. Also note the unexpectedly low performance of GBT for both feature sets.

We also investigated the performance of k-NN as the function of k in Fig. 10. As we observed, in all cases, $k = 2$ performs best with slow degradation and finally slight increase towards the global average as prediction.

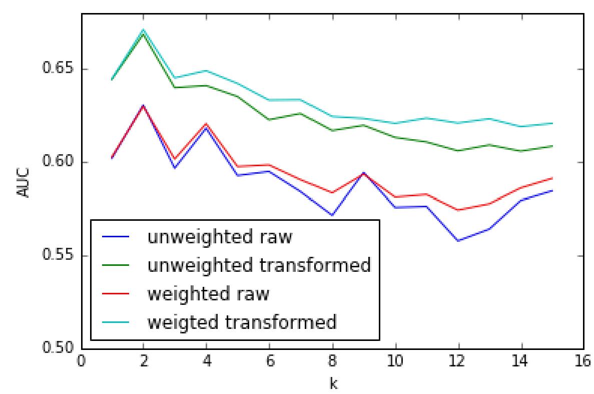

Figure 10: Performance of DTW based k-NN as function of k.

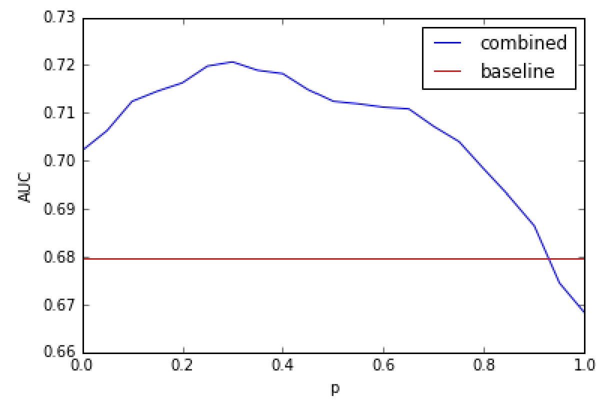

Figure 11: AUC for the combined predictions $(1 - p) \cdot$ 2-NN DTW $+ p \cdot$ SVR on time series features. Horizontal line is the AUC for the prediction SVR on scrap rates.

We combined the best time series feature based model prediction, SVR on transformed features, with the 2-NN time series similarity based prediction. As a baseline, we considered SVR on scrap rates, since it had the best performance among the scrap rate based models. The result in Fig. 11 shows that the combined models outperform the baseline, i.e. we are able to predict the undesirable increase of scrap rate just looking at the process parameters with better accuracy than models using information about past scrap rates. Our model hence gives warnings for potentially unexpected scrap rate increase to the operators.

Finally we investigate how prediction quality depends on the average scrap rate around the time of the prediction. Recall from Fig. 7 that we have three periods, the first with normal, the second with increased, and the last with very high scrap rates. The best prediction by using the combined models is depicted in Fig. 12. We observe that the warnings are present in all phases, and thus our solution may be meaningful under different circumstances. By our careful selection of removed features that characterize the production period such as outside air temperature, we were also able to avoid overfitting to the periods with high scrap rate.

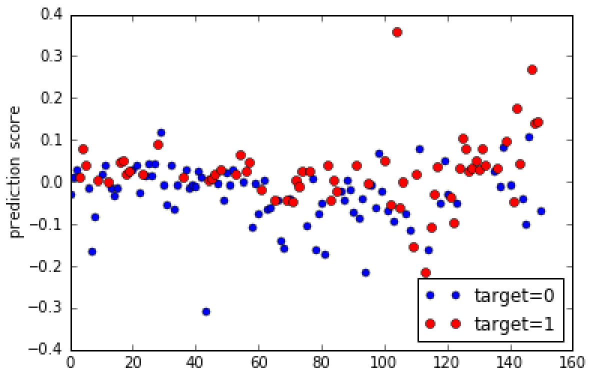

Figure 12: Prediction scores of the best combination as a series of the 164 production cycles. Actual positive instances are red.

7. RELATED RESULTS

Data driven methods have a wide range of applications in manufacturing, both in the product development phase and during manufacturing. Predictive methods can be used for quality improvement in various ways [9] by describing product and process quality, predicting quality both for design and during manufacturing, distinguishing machine or product failure patterns, and optimizing process parameters especially in the development phase.

In this paper, we used state-of-the-art classification techniques [22]. The use of DTW [10] is common in several fields including speech, gesture and shape recognition. The efficient calculation of DTW is studied in [8,14].

Surprisingly, time series classification methods known well for over a decade in the machine learning community have apparently not yet reached the manufacturing application domain. A survey from 2011 [9] on data mining applications for quality improvement in manufacturing industry has no mention of time series classification. Next we review results concerning data analysis of the molding process. While the present paper considers the transfer molding process, most related results are on injection molding technologies [15]. These work with raw material having different characteristics than transfer molding, however the parameters which should be controlled carefully are basically the same.

While our focus is predicting quality in future production batches, related research focuses only on predicting individual product quality. In our case, as the final product is ultrasonar scanned, predicting scrap after production is meaningless. In other applications, however, researchers may build models from actual or simulated production data for the purpose of optimizing process parameters.

Typical results for analyzing molding process data predict individual product quality based on rather limited feature sets. For example, [13] uses only mold flow rate, injection pressure, mold temperatures, and melt temperatures as input to artificial neural networks. They predict the quality of simulated injection molding only. In a paper using neural networks for injection molding scrap prediction [3] that lists several other similar results, all cited papers use 4-6 features only.

In case of molding, parameter optimization is made by making experiments with different parameter settings, and then interpolating the value of the quality index. In [1,16,17] after selecting some features, neural networks and genetic algorithms are used to predict the quality of a product manufactured with a given setting, and thus finding the optimal process parameters.

Another important application of data analysis in molding is online quality monitoring. [18] presents a real case application, where pressure sensor data for few hundred products was collected from manufacturing with different parameters, and the data is classified as good or scrap with high accuracy. In [12] an in-process data is inspected to classify six different failure modes after the product left the machine using neural networks and SVM. In [24] a method is presented for on-line quality prediction in transfer molding. First, principal component analysis is used to determine few uncorrelated features in strong connection with the quality index. In order to take the mold cycles of the past into consideration, an ARX model is used for predicting the quality. The method is tested with several parameter settings on a production period.

We found only two papers, both mentioned above, that create time series based on sensor data of molding processes, however even these papers apply standard classification by using points of the time series as regular numeric features. In [12], temperature and pressure time series are created, however due to the large input size, they only kept six process variables: cycle time, metering time, injection time, barrel temperature at one stage, cushion, and injection velocity. Finally in [18], pressure time series were collected by using two sensors. They recorded 237 pressure cycles, out of which 85 led to producing faulty pieces. They report accuracy for using Naive Bayes, decision trees, SVM and nearest neighbors after clustering. However they use no time series specific methods, rather just use pressure values as numeric features for the classifiers.

8. CONCLUSION

In this paper we demonstrated how time series classification can aid manufacturing processes by issuing warnings for anticipated increase of scrap rates. On a 3-month molding manufacturing data, we simulated a warning system for scrap rate increase by using different classification methods, and evaluated the predictions. As result, we found dynamic time warping based nearest neighbor and time series based support vector regression as good performing methods that also combine well in a classifier ensemble. The crux of our results lie in feature engineering: we process hierarchical data, in which sensor time series of individual products are arranged in a higher level time series of the production lifetime. The presented results can serve as a solid basis for designing and deploying a warning system that could work in real time. We plan to release a slightly blurred version of the data for the reproducibility of our experiments.

9. REFERENCES

[1] I. Ali and Y.-T. Chen. Design quality and robustness with neural networks. *IEEE Transactions on Neural Networks*, 10(6):1518–1527, 1999.

[2] C. Bahlmann, B. Haasdonk, and H. Burkhardt. Online handwriting recognition with support vector machines-a kernel approach. In *Frontiers in handwriting recognition, 2002. proceedings. eighth international workshop on*, pages 49–54. IEEE, 2002.

[3] W.-C. Chen, P.-H. Tai, M.-W. Wang, W.-J. Deng, and C.-T. Chen. A neural network-based approach for dynamic quality prediction in a plastic injection molding process. *Expert systems with Applications*, 35(3):843–849, 2008.

[4] B. Daróczy, P. Vaderna, and A. Benczúr. Machine learning based session drop prediction in lte networks and it s son aspects. In *Proceedings of IWSON at IEEE 81st Vehicular Technology Conference (VTC Spring), 2015, Glasgow, Scotland, UK*, 2015.

[5] J. Fogarty, R. S. Baker, and S. E. Hudson. Case studies in the use of roc curve analysis for sensor-based estimates in human computer interaction. In *Proceedings of Graphics Interface 2005*, GI '05, pages 129–136, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2005. Canadian Human-Computer Communications Society.

[6] J. H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232, 2001.

[7] S. Gudmundsson, T. P. Runarsson, and S. Sigurdsson. Support vector machines and dynamic time warping for time series. In *Neural Networks, 2008. IJCNN 2008.(IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*, pages 2772–2776. IEEE, 2008.

[8] E. J. Keogh and M. J. Pazzani. Scaling up dynamic time warping for datamining applications. In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '00, pages 285–289, New York, NY, USA, 2000. ACM.

[9] G. Köksal, İ. Batmaz, and M. C. Testik. A review of data mining applications for quality improvement in manufacturing industry. *Expert systems with Applications*, 38(10):13448–13467, 2011.

[10] J. B. Kruskall and M. Liberman. The symmetric time warping algorithm: From continuous to discrete. *Time Warps, String Edits and Macromolecules: The Theory and Practice of String Comparison*, 1983.

[11] A. Mangal and N. Kumar. Using big data to enhance the bosch production line performance: A kaggle challenge. *arXiv preprint arXiv:1701.00705*, 2016.

[12] 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, 2005.

[13] B. Sadeghi. A BP-neural network predictor model for plastic injection molding process. *Journal of materials processing technology*, 103(3):411–416, 2000.

[14] S. Salvador and P. Chan. Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, 11(5):561–580, 2007.

[15] C. M. Seaman, A. A. Desrochers, and G. F. List. Multiobjective optimization of a plastic injection molding process. *IEEE Transactions on Control Systems Technology*, 2(3):157–168, Sep 1994.

[16] C. Shen, L. Wang, and Q. Li. Optimization of injection molding process parameters using combination of artificial neural network and genetic algorithm method. *Journal of Materials Processing Technology*, 183(2):412–418, 2007.

[17] H. Shi, Y. Gao, and X. Wang. Optimization of injection molding process parameters using integrated artificial neural network model and expected improvement function method. *The International Journal of Advanced Manufacturing Technology*, 48(9-12):955–962, 2010.

[18] A. Tellaeche and R. Arana. Machine learning algorithms for quality control in plastic molding industry. In *2013 IEEE 18th Conference on Emerging Technologies & Factory Automation (ETFA)*, pages 1–4. IEEE, 2013.

[19] H. T. Tran, M. H. Shirangi, X. Pang, and A. A. Volinsky. Temperature, moisture and mode-mixity effects on copper leadframe/emc interfacial fracture toughness. *International Journal of Fracture*, 185(1-2):115–127, 2014.

[20] V. Vapnik. The support vector method of function estimation. In *Nonlinear Modeling*, pages 55–85. Springer, 1998.

[21] S. Wang, J. Wan, D. Li, and C. Zhang. Implementing smart factory of industrie 4.0: an outlook. *International Journal of Distributed Sensor Networks*, 2016.

[22] X. Xi, E. Keogh, C. Shelton, L. Wei, and C. A. Ratanamahatana. Fast time series classification using numerosity reduction. In *Proceedings of the 23rd international conference on Machine learning*, pages 1033–1040. ACM, 2006.

[23] Xie and Sitaraman. Investigation of interfacial delamination of a copper-epoxy interface under monotonic and cyclic loading: experimental characterization. *IEEE Transactions on Advanced Packaging*, 26(4):447–452, 2003.

[24] S. Zhang, R. Dubay, and M. Charest. A principal component analysis model-based predictive controller for controlling part warpage in plastic injection molding. *Expert Systems with Applications*, 42(6):2919–2927, 2015.