Augmentation of Body-in-White Dimensional Quality Systems through Artificial Intelligence

Carlos A. Escobar
Research and Development
General Motors
Warren, USA
ORCID 0000-0002-7234-8175

Debejyo Chakraborty
Research and Development
General Motors
Warren, USA
ORCID 0000-0002-4816-5467

Jorge Arinez
Research and Development
General Motors
Warren, USA
ORCID 0000-0003-4893-6935

Ruben Morales-Menendez
Escuela de Ingenieria y Ciencias
Tecnologico de Monterrey
Monterrey, NL, Mexico
ORCID 0000-0003-0498-1566

2021 IEEE International Conference on Big Data (Big Data) | 978-1-6654-3902-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/BigData52589.2021.9671610

***Abstract*—Reducing the dimensional variability of the body-in-white (BIW) in automotive manufacturing is perhaps the most difficult quality control problem due to complex interdependencies amongst the multiple assembly stations that a BIW must pass through in a bodyshop. As increasing quantities of dimensional data are generated in factories, manufacturers face the challenge and opportunity to derive value from the data by enabling advanced quality control methods that can realize greater dimensional stability. As the BIW moves through the bodyshop, dimensional deviations propagate and amplify to downstream stations affecting final vehicle fit-and-finish and visible quality aesthetics potentially influencing a customers' purchase decision. Current BIW quality approaches rely on univariate statistical process control (SPC) charts. With the large amounts of complex data produced, such charts often fail to detect quality patterns that may exist in hyper-dimensional spaces. As a stop-gap measure, manufacturers attempt to remediate quality issues by assigning operators in final vehicle assembly to visually identify and manually fix apparent deviations. This paper illustrates the application of artificial intelligence (AI) to develop a real-time monitoring system that seeks to predict and detect early dimensional quality issues and eliminate the need for costly downstream corrective actions. Moreover, beyond early detection and prediction, the proposed system also facilitates diagnosis of root causes and understanding the true nature of quality issues.**

***Index Terms*—body-in-white, dimensional data, machine Learning, quality control, smart manufacturing, human intelligence augmentation**

I. Introduction

In automotive manufacturing, the assembly of the Body-in-white (BIW) that occurs in the bodyshop of a vehicle assembly plant, one of the most challenging quality control problems in this type of multistage assembly system is that of dimensional variability reduction. Dimensional deviations from nominal are propagated and amplified from upstream to downstream stations. Consequently, dimensional errors accumulate leading to non-trivial quality issues in the final vehicle assembly line of the plant – at this late stage in a vehicle's manufacture, such issues are much more difficult and expensive to repair. These dimensional discrepancies in the final vehicle's build may be visible to the customer and can potentially negatively affect a customers' purchase decision. To improve overall vehicle build quality, manufacturers seek to develop advanced quality control methods as a means to ensure dimensional stability. This work describes the application of artificial intelligence (AI) to develop a real-time monitoring system that provides early detection of dimensional quality issues and augments the role of human intelligence to expedite trouble-shooting processes.

To improve the quality and optimization of processes in *predictive manufacturing systems* (*PMS*), one of the most promising areas of AI is machine learning (ML) [1]. ML techniques have the ability to identify implicit relationships in data, especially in large volumes that change over time [2]. In manufacturing, process optimization [3], fault detection [4] and predictive maintenance [5], are among the prominent areas of implementation of ML.

Different ML tools can provide different benefits. Techniques like *random forest* (RF) in milling operations can identify the relationship of cutting force, vibration, and acoustic emission signals with that of cutting tool wear. In addition, speech recognition frameworks have been exploited successfully to monitor tool wear [6], [7]. This particular work utilized a continuous hidden Markov model classifier where the feature vectors consisted of the Mel frequency cepstrum coefficients of the vibration signal. These results allowed the development of supervisory control systems [8]. Another example using RF is of a real-time monitoring manufacturing system to detect faults based on sensor data combining RF classifiers and outlier detection [9].

Gradient boosted trees are algorithms that have shown promising results in manufacturing to identify nonlinear patterns and anomalies in data sets for real-time prediction of weld quality in metal active gas welding processes [10], [11]. There are also decision-making tools based on this method to support operators to classify defects in the manufacturing systems [12].

As an uncertainty modeling tool, the *Bayesian network* (BN) approach has been extensively applied to fixturing fault diagnosis. A BN is a probabilistic graphical model representing a set of random variables and their conditional dependencies via a directed acyclic graph. Several applications of BNs in assembly systems were evaluated by [13]. These applications include monitoring/diagnosis in the multistage cap alignment process [14], fixture failure diagnosis [15], a systematic approach for process fault diagnosis based on *BN* considering incomplete data and varying noise levels. Additional application examples of *ML* in manufacturing systems include topological data analysis [16], deep learning [17], and genetic algorithms to evaluate form tolerances [18].

To monitor final BIW quality, *online optical coordination measurement machines* (OCMMs) are widely used in the automotive industry. They are frequently installed at the end of each assembly line and can use more than 100 optical laser sensors to measure key measurement points (MP) set at the final BIW. The effectiveness of using control charts largely depends on the correct recognition for different kinds of variation patterns, such as control chart pattern (CCP) [19]. The common CCP recognition approaches can be subsumed under just two broad categories: run-rule-based *expert systems* (ES) [20], and ML methods like support vector machines (SVM) [21] and artificial neural network (ANN) [22]. ES contains information explicitly, i.e. they can be modified/updated easily. However, there is no one-to-one correspondence between failures and a running rule; this requires engineering experience. SVM shows generalization in small problems, but it is not easy to select the right kernel functions and optimize their parameters. ANN exhibits nonparametric, nonlinear adaptive learning, however, it needs large volumes of training data. Long short-term memory neural network (LSTMNN) are a recurrent neural network (RNN) structure to model temporal sequences and their long-term dependencies. LSTMNN could be well suited for dealing with serially correlated or auto-correlated data to recognize variation patterns from BIW OCMM datasets. LSTMNN may also effectively eliminate the confusion between different patterns, and accurately identify a variety of abnormal variation patterns in BIW OCMM data. The robustness of LSTMNN variation recognition and may also help engineers to improve the efficiency/accuracy of fault diagnoses based on the fault cause database [23].

In our BIW problem, we chose to exploit the simplicity of a shallow neural network and focused the computation on selecting the simplest network that would yield the highest accuracy.

During the BIW manufacturing process, the underbody structure is inspected by a vision system where the Cartesian coordinates in three dimensional space of the salient underbody dimensional points are recorded as $\boldsymbol{u} \in \mathbb{R}^3$. Here, $\mathbb{R}^3$ denotes a set of three-dimensional real numbers. The observed coordinates are subsequently compared to nominal values $\tilde{\boldsymbol{u}}$ to obtain the deviations $\boldsymbol{\delta} \in \mathbb{R}^3$ from nominal. The deviation of the ith ($i \in \mathbb{N}$, set of natural numbers) point is defined as $\boldsymbol{\delta}_{\boldsymbol{u}_i} = \boldsymbol{u}_i - \tilde{\boldsymbol{u}}_i$. These deviations are monitored using univariate control charts, Figure 1. The underbody and the parts are staged in a holding fixture as shown in Figure 2. The parts are clamped according to geometric dimensioning and tolerancing requirements, manually compensating for the dimensional error through actions like *shim moves/alterations*. This underbody assembly is then welded together with the body sides and roof in a framing station yielding the complete BIW shown in Figure 3. The BIW then undergoes a vision inspection similar to the underbody previously, recording corresponding dimensional points as $\boldsymbol{f} \in \mathbb{R}^3$, generating the univariate control charts as shown in Figure 4. The observed

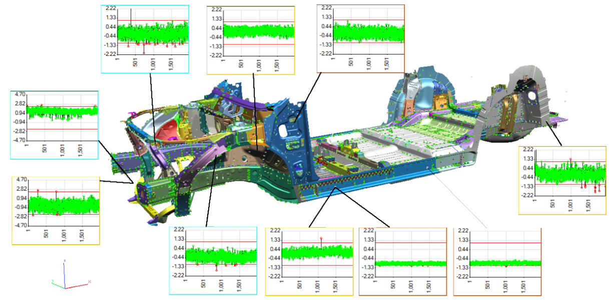


Fig. 1. Underbody control charts.

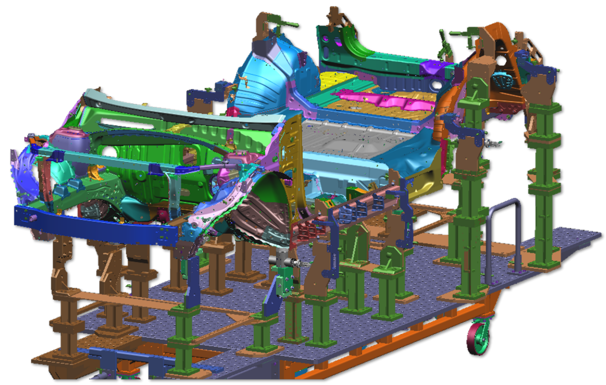


Fig. 2. Underbody mounted on a holding fixture.

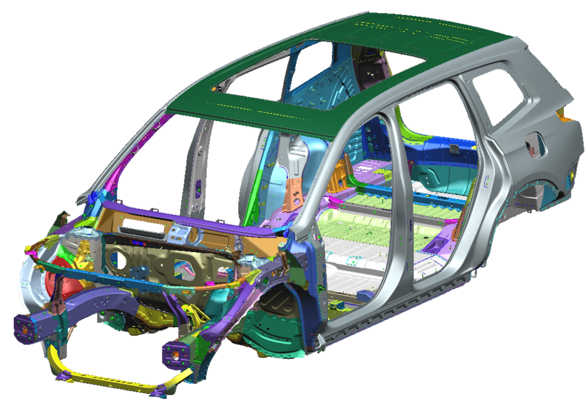


Fig. 3. Complete assembled BIW.

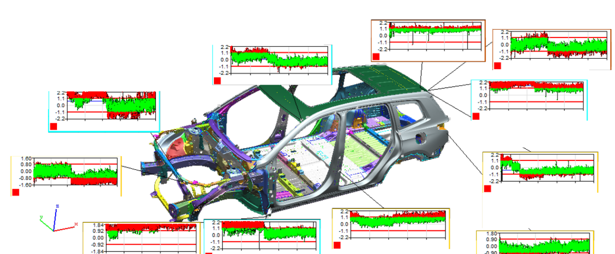


Fig. 4. Framing station control charts.

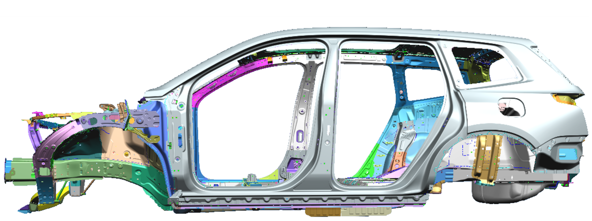
points are then compared against nominal points $\tilde{\boldsymbol{f}} \in \mathbb{R}^3$ to obtain the deviations of any jth point $\boldsymbol{\delta}_{\boldsymbol{f}_j} = \boldsymbol{f}_j - \tilde{\boldsymbol{f}}_j$. The existing body is then manually altered to minimize the dimensional discrepancy. Process changes through manual adjustments, additional shim moves/alterations at the underbody staging, are then made to minimize the discrepancy in the next body. These corrective actions are based on empirical knowledge of the dimensional engineers.

Univariate monitoring, though somewhat effective, is incapable of detecting multi-dimensional concerns. Consequently, framing station quality issues are observed (Figure 4) though all the underbody points were *under control*, or within limits as shown in Figure 1.

In this study, artificial intelligence is applied to augment the current quality control system. All the underbody dimensional data is used in a multi-variate approach to develop a predictive system. This system is aimed at detecting quality patterns that exist in a hyperdimensional space. The objective is to develop *smart* shim moves, i.e., modifications to the system input parameters that, according to the predictive system, reduce deviations. The problem is formulated as a regression.

The hypothesis here is that the underbody dimensional deviations hold multivariate linear/non-linear relationships with framing station deviations. To test this hypothesis, machine learning techniques were applied to the underbody coordinates $\boldsymbol{u}$ to predict framing station deviations $\boldsymbol{\delta}_{\boldsymbol{f}}$ in final assembly. If the hypothesis held true, learned relationships were then used to augment human intelligence to expedite trouble-shooting processes throughout *smart shim moves*. A representative BIW process is used as a case study, Figure 5.

(a) Completed underbody assembly



(b) Completed BIW following framing station.

Fig. 5. A representative body assembly depicting the two main process steps to marry underbody with uppers (body sides and roof).

In the next section, Section II, an overview of the solution is provided. Subsequently, the modeling results is presented through a case study in Section III. Finally, Section IV presents the conclusions.

II. Solution Overview

The solution has two components aimed at augmenting control charts with human intelligence. Figure 6 shows the solution framework.

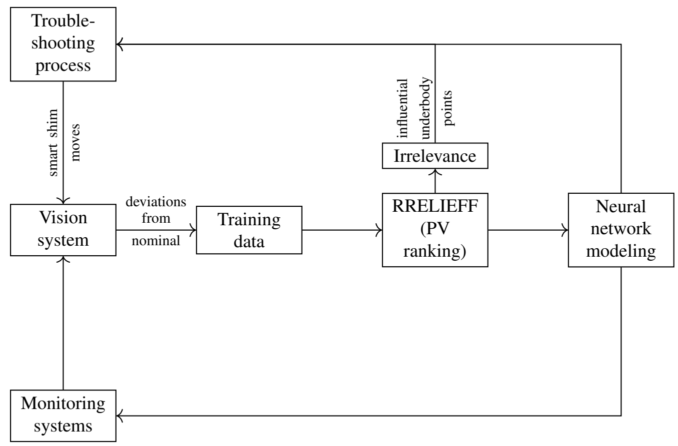


Fig. 6. Solution framework

The first component is the monitoring system. An artificial neural network algorithm is applied to learn relationships between the underbody and the framing station deviations. One neural network structure is developed for each framing station coordinate as shown in Figure 7. Underbody data from each vehicle is run through the trained neural network, in real-time, to predict framing station dimensional quality issues, i.e., deviations that would exceed predefined tolerances.

The other component is a troubleshooting process. If quality issues are predicted, the trained neural network model is also used to perform virtual sensitivity analyses to guide the physical adjustments, *smart shim moves*. Moreover, the most influential underbody points are also identified using the RReliefF algorithm [24]. This information helps to prioritize the dimensional engineering efforts by identifying the driving points associated to the overall quality of the frame.

A. Modeling Details

The inputs are the individual dimensions of the m underbody deviations denoted by set $\mathbb{D} = \{\mathfrak{d}_k | k = 1, \ldots, m\}$ where $\mathfrak{d}_k \in \mathbb{R}$ is any one dimension (x-, y- or z-) of an underbody deviation point $\boldsymbol{\delta}_{\boldsymbol{u}_i}$. The goal of the model is to predict the framing station deviations in $\mathbb{F} = \{\mathfrak{f}_l | l = 1, \ldots, n\}$ where each element $\mathfrak{f}_l \in \mathbb{R}$ is a dimension of $\boldsymbol{\delta}_{\boldsymbol{f}_j}$. In order to minimize the modeling effort, RReliefF algorithm is used to ascertain the most relevant elements of $\mathbb{D}$ that could predict any $\mathfrak{f}_l$. The RReliefF algorithm provides a weight $w \in \mathbb{R}$ of relevance for each $\mathfrak{d}_k$ in predicting every $\mathfrak{f}_l$. The weights could be arranged in a matrix $W \in \mathbb{R}^{m \times n}$. The relevant $\mathfrak{d}$ are then identified by thresholding these weights. If any weight $w_{k,l}$ exceeds a threshold τ then $\mathfrak{d}_k$ is relevant in predicting $\mathfrak{f}_l$.

The weights are also assigned ranks $r \in \mathbb{N}$. A rank of $r_{k,l} = 1$ is assigned to the highest among weights $w_{\cdot,l}$, implying that $\mathfrak{d}_k$ is the most relevant dimension of an underbody

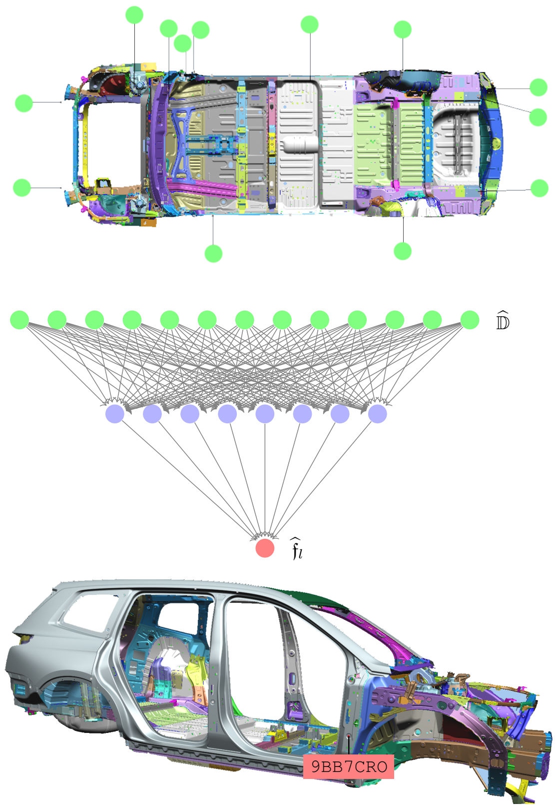


Fig. 7. Mapping tolerance in underbody points to a framing station point `9BB7CRO`.

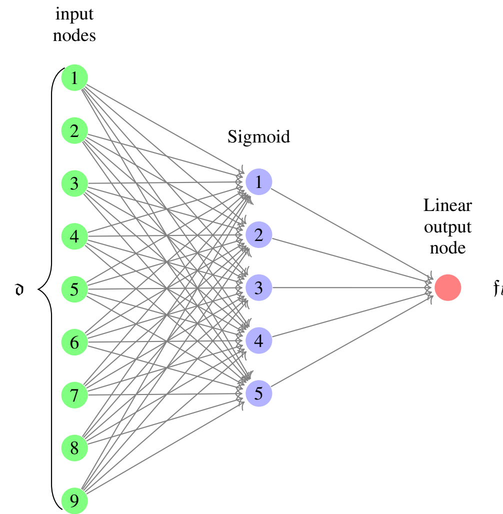


Fig. 8. Architecture of network $a = 9$.

deviation that can predict $\mathfrak{f}_l$. The ranks are also arranged in $R \in \mathbb{R}^{m \times n}$, corresponding to the weights.

Several different feed-forward neural network (FFNN) [25] models were created, for predicting each $\mathfrak{f}_l$ by varying the training set. The first network, $a = 1$ was trained on only one input $\mathfrak{d}_k$ for which $r_{k,l} = 1$. The second network, $a = 2$ included all $\mathfrak{d}_k$ such that $r_{k,l} \le 2$, and so on. The last model, model $a = a_l$ included all the $a_l = \max(r_{\cdot,l})$ elements of $\mathbb{D}$ that were deemed relevant in predicting $\mathfrak{f}_l$.

The chosen FFNN contained three layers: input, hidden, and output, as depicted in Figure 8. In such a network, information moves only in forward direction. It enters the network through the input neurons at the first layer. A mathematical process is then developed [26] at the next layer, the hidden layer, using activation functions [27]. A sigmoid activation function was chosen in this architecture. The outcome of this layer is then weighted and fed to the final layer, which provides the outcome. The number of nodes in the hidden layer is empirically determined [28]. According to [29], the optimal size of the hidden layer is usually between the size of the input and size of the output layers. For predicting any $\mathfrak{f}_l$ we picked $\lceil \frac{a_l+1}{2} \rceil$ to determine the number of nodes in the hidden layer.

The output layer had one output node with a linear transfer function. The Levenberg-Marquardt algorithm [30] was used to trained the network. Using a hold-out validation scheme (70% training, 15% validation, 15% testing), the training continued until the validation error failed to decrease for six iterations. The trained network was then used to predict the outcome of the test set.

A network score

$$\mathcal{S}_{a_l} \triangleq \sqrt{b_\rho \frac{|\rho_l|^3}{|\rho_l| + \rho_l} + b_\varepsilon \left(\frac{\varepsilon_l - \min(\varepsilon_l)}{\max(\varepsilon_l) - \min(\varepsilon_l)} \right)^2 + b_a \left(\frac{a-1}{a_l - 1} \right)^2} \tag{1}$$

to compare the models was computed using Pearson's correlation coefficient between predicted and observed values ($\rho \in [-1, 1]$), mean squared error (MSE) $\varepsilon \in \mathbb{R}$, and a_l. This scoring mechanism rewards high ρ (while severely penalizing for $\rho < 0$), low MSE and penalizes for large model complexity [31]. In this score, the weights assign importance to each metric and assume the following values

$$\begin{aligned} b_\rho &= 2, \\ b_\varepsilon &= 1, \\ b_{a_l} &= 0.1. \end{aligned}$$

B. *Influential Points Identification*

The prediction of framing station deviations can provide further engineering understanding with a metric that determines the significance of the influence of an underbody point on a framing station point. Such a metric is established through the calculation of irrelevance whereby the smaller the irrelevance, the more relevant/influential a point is.

Recall that every deviation $\boldsymbol{\delta}$ is composed of three orthogonal components along the x-, y-, z- directions $\delta_x, \delta_y, \delta_z \in \mathbb{R}$. Thus, three elements from $\mathbb{D}$ would make up $\boldsymbol{\delta}_{\boldsymbol{u}_i}$ and three elements from $\mathbb{F}$ would make up $\boldsymbol{\delta}_{\boldsymbol{f}_j}$. With a latent look up table function Δ of $i \to k$ and $j \to l$ mapping, any $\boldsymbol{\delta}$ could be recomposed from $\mathbb{D}$ and $\mathbb{F}$, respectively. Let us say, in the lack of a closed form mathematical formulation, that Δ is a variant of Kronecker delta function such that the index k could pick all the dimensions of $\boldsymbol{u}_i$ and l could pick all the dimensions of $\boldsymbol{f}_j$.

The overall irrelevance γ of an underbody point $\boldsymbol{u}_i$ to any framing station point $\boldsymbol{f}_j$ is defined as the average of 2-norm of 3-dimensional ranks

$$\gamma_{i,j} \triangleq \left\lfloor \mathcal{C}_1 \frac{1}{3} \sum_k \sqrt{\sum_l R(k\Delta_{i,k}, l\Delta_{j,l})^2} + \mathcal{C}_0 \right\rfloor , \qquad (2)$$

Here $\mathcal{C}_1, \mathcal{C}_0 \in \mathbb{R}$ are constants and $\gamma \in \mathbb{N} \cap [1, 256]$. In this dataset, not all x-, y-, z- directions were observed for every point since the existing method measured univariate dimensions (Figures 1 and 4). The missing dimensions were filled in as an average of the existing dimensions. For example if the x- dimension u_{i_x} of any point $\boldsymbol{u}_i$ was the missing dimension, then to calculate irrelevance of that point in estimating $\boldsymbol{f}_j$ would be achieved by generating the rank of u_{i_x} as

$$r_{i_\mathrm{x}}|j = \frac{1}{2}(r_{i_\mathrm{y}}|j + r_{i_\mathrm{z}}|j).$$

Note that while $r_{i_\mathrm{x}}|j$ is not in R, $r_{i_\mathrm{y}}|j$ and $r_{i_\mathrm{z}}|j$ are in R. However, if two of the dimensions were missing, the available third dimension value was simply replicated.

III. Case Study — Modeling Results

A data set derived from the representative BIW process shown previously in Figure 5. This data set comprised 9,099 samples with each containing 118 underbody station dimensions, i.e. $|\mathbb{D}| = 118$, and 79 framing station dimensions, i.e $|\mathbb{F}| = 79$. The data was divided into 5,459 training, 1,820 validation, and 1,820 testing groups. The modeling approach using this data is first discussed in Section III-A. Subsequently, in Section III-B a physical orientation of the influential points is established based on their irrelavance metric.

A. Modeling

Using the RReliefF algorithm, ranks were assigned to all the $\mathfrak{d}$ for every $\mathfrak{f}$ as shown in Figure 9. As we would soon see that

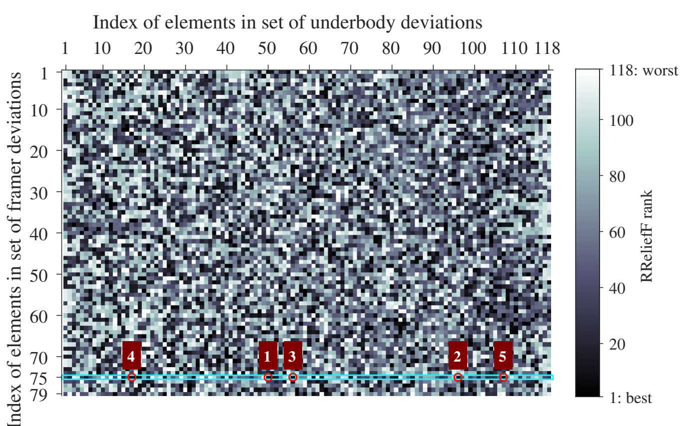


Fig. 9. RReliefF ranking of all framing station variables for each underbody variable, highlighting the best predicted framing station variable $\mathfrak{f}_{75}$ and its top 5 underbody points, $\mathfrak{d}_{50}, \mathfrak{d}_{96}, \mathfrak{d}_{56}, \mathfrak{d}_{17}, \mathfrak{d}_{107}$.

our prediction was the best for the y-dimension of 2632COF $\mathfrak{f}_{75}$, let us use that for illustrating an example[1]. Figure 10 shows the actual weights that were generated using RReliefF for this framing station dimension. A weight threshold of

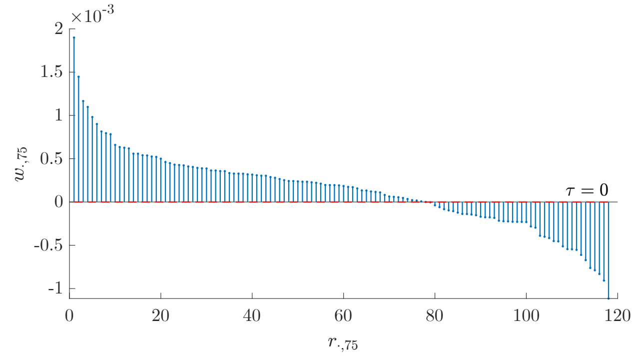
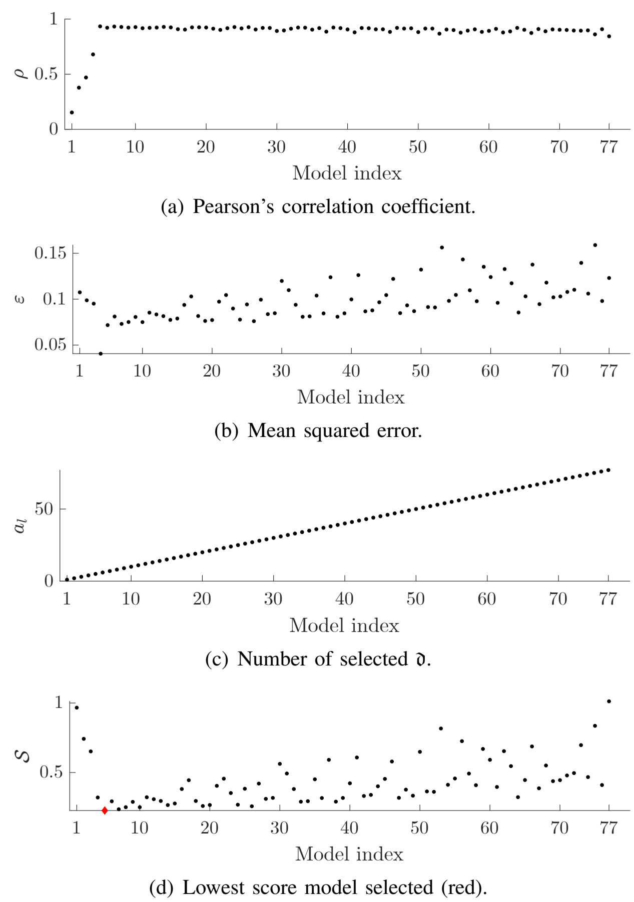


Fig. 10. Rank versus weights for $l = 75$.

$\tau = 0$ was used to select the 77 relevant underbody points. All 77 models were created and evaluated with the validation group to calculate the score according to Equation (1). The modeling results are provided in Figure 11. The fifth model



(a) Pearson's correlation coefficient.

(b) Mean squared error.

(c) Number of selected $\mathfrak{d}$.

(d) Lowest score model selected (red).

Fig. 11. Modeling result.

(i.e. the one using five inputs) had the best (lowest) score of

[1]the labels are obfuscated

$\mathcal{S} = 0.2266$. Thus, the top five ($r_{\cdot,75} \leq 5$) ranking underbody dimensions for this framing station dimension $\mathfrak{f}_{75}$ were used to construct the final model. This model contained three hidden nodes. These underbody dimensions were identified to be $\mathfrak{d}_{50}, \mathfrak{d}_{96}, \mathfrak{d}_{56}, \mathfrak{d}_{17}, \mathfrak{d}_{107}$, in the order of increasing rank as shown in Figure 9. Evident from the high Pearson's coefficient and low MSE, $\mathfrak{f}_{75}$ was highly predictable. Figure 12 compares

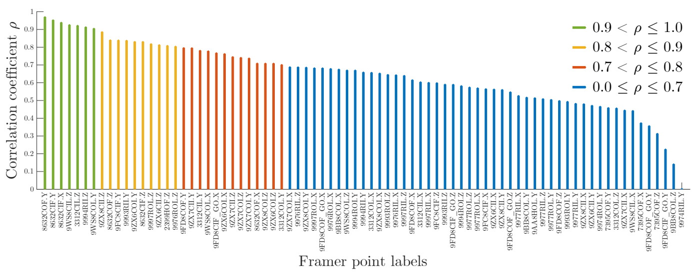


Fig. 12. Prediction results of each $\mathfrak{f}_l$ using the test set.

the Pearson's correlation coefficient of all the frame dimension in $\mathbb{F}$ and $\mathfrak{f}_{75}$ was indeed the one with the highest correlation.

As for the entire population of the framing station dimensions, seven (out of 79) framing station coordinates exhibited excellent predictability ($\rho \geq 0.90$), 12 coordinates exhibited very good predictability ($0.90 > \rho \geq 0.80$), 11 good predictability ($0.80 > \rho \geq 0.70$), and 49 fair ($\rho < 0.70$).

Once this model was finalized, the test group was used to test the model. The neural network model generalized well to the test group with $\rho = 0.9457$, and $\varepsilon = 0.0751$. A plot of the observed and predicted values for this dimension as shown in Figure 13 demonstrates that the model was sensitive

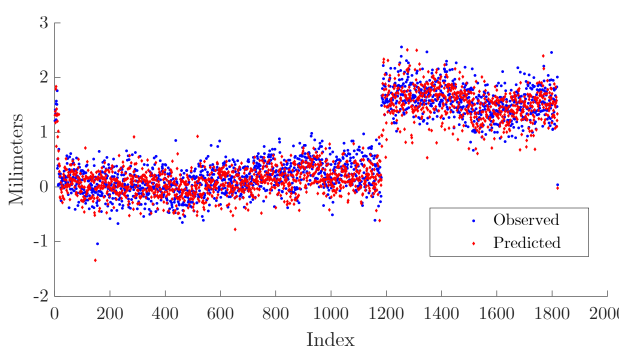


Fig. 13. Observed vs predicted values in the test set.

to mean shift, at around data index 1,200. The error histogram in Fig. 14 shows a nearly zero mean error with most of the error within ± 0.5 mm, making the model prediction reliable.

B. Physical Interpretation and Relevance

The significance of the underbody point in predicting a framing station point was measured using irrelevance as defined in Equation (2). The smaller the irrelevance, the larger the relevance. For framing station point 8632COF, of which the y-direction was $\mathfrak{f}_{75}$, the irrelevance of the 55 best underbody points are shown in Figure 15. The most relevant

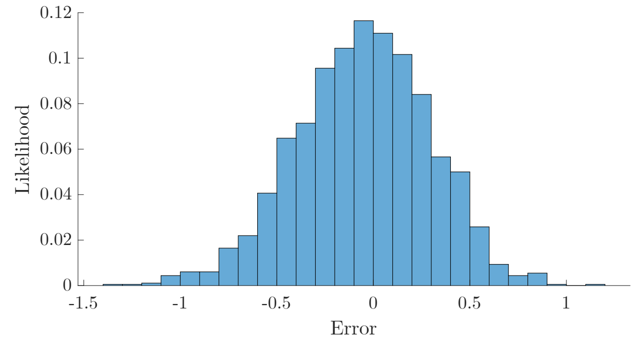


Fig. 14. Histogram of errors on the test set.

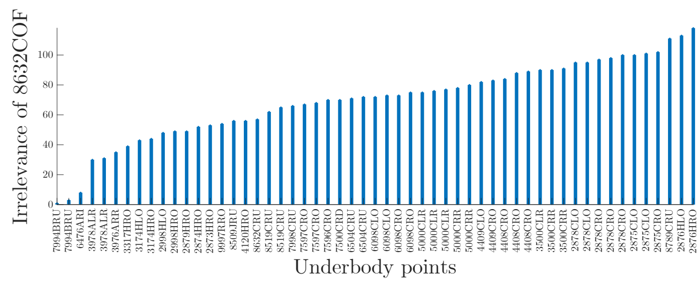


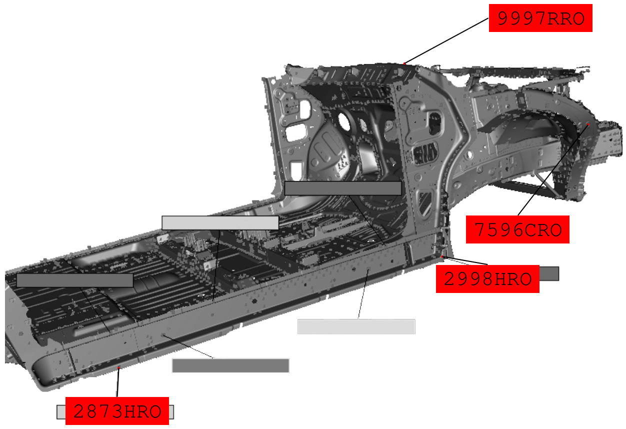
Fig. 15. Relevance of the selected underbody points in predicting the highest correlation framing station point.

underbody point, for predicting 8632COF framing station point, was 7994BRU. Such information is invaluable to the engineers. In fact, by aggregating the irrelevance over all the framing station points, one could easily see from Figure 16(b) which underbody points are most crucial to closely control to minimize any error in the framing station. The size of the marker is proportional to the relevance while the color indicated the irrelevance. The biggest marker has the smallest irrelevance. The top five most relevant points are marked with black font.

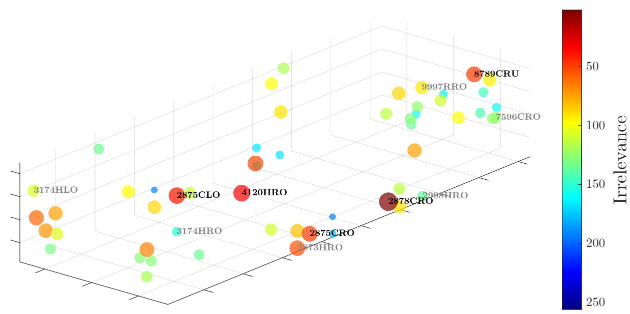
Figure 16(a) provides a visual perspective which includes some labelled reference underbody points. The gray font in Figure 16(b) indicates some of these reference underbody points.

IV. Conclusion

Vision-based dimensional quality measurement systems generate a large quantity of discrete, three-dimensional coordinate data that capture deviations from nominal and represent an opportunity to increase understanding and potentially improve prediction and thereby control to reduce variation during the BIW assembly process. The conventional monitoring of coordinate dimensions in a serial manner does not readily yield the identification of anomalies that may be occurring in three-dimensional space across multiple assembly stations. Moreover, understanding the (non-)linear relationships between the underbody station and framing station deviations is made more challenging given the multivariate nature of the measurement data. Historically, maintaining dimensional quality in such complex body assembly systems has had to



(a) Selective physical location of some underbody points.



(b) Overall irrelevance.

Fig. 16. Locations and influence of underbody points.

rely significantly on domain experience and inefficient, trial-and-error adjustment approaches.

In this study, we demonstrated the applicability of a neural network model to early prediction of potential BIW quality issues. The model showed the use of underbody measurement data to predict potential concerning deviations arising downstream at the framing station. Such a model establishes the link between dimensional trends appearing at the underbody station and their effect on final BIW quality at the framing station. There are numerous benefits that such a linking model affords, chief among them is the ability to anticipate issues and to rapidly resolve the root cause of deviations should they occur.

Finally, the information derived from this machine learning approach was also used to explain which underbody points are important to closely monitor and those which are insignificant and do not merit attention. This distinction alone is important since engineering problem solving effort can be more effectively focused on the dimensions where issues originate. As such, this invaluable insight distills a large set of possible points into the critical ones that effect final product quality. However, these inferences should be used with caution as there may be other unaccounted for sources of variation that could render the process unstable. In these circumstances, periodic relearning is advised to ensure the model stays relevant as processes evolve. An important area for further research work is to understand the relevant process dynamics to appropriately implement relearning schemes and their scheduling.

References

[1] G. Köksal, I. Batmaz, and M. C. Testik, "A Review of Data Mining Applications for Quality Improvement in Manufacturing Industry," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 13 448–13 467, 2011.

[2] T. Wuest, D. Weimer, C. Irgens, and K.-D. Thoben, "Machine Learning in Manufacturing: Advantages, Challenges, and Applications," *Prod. Manuf. Res.*, vol. 4, no. 1, pp. 23–45, 2016.

[3] D. Kwak and K. Kim, "A Data Mining Approach Considering Missing Values for the Optimization of Semiconductor-Manufacturing Processes," *Expert Syst. Appl.*, vol. 39, no. 3, pp. 2590–2596, 2012.

[4] D. Kim, P. Kang, S. Cho, H. Lee, and S. Doh, "Machine Learning-based Novelty Detection for Faulty Wafer Detection in Semiconductor Manufacturing," *Expert Syst. Appl.*, vol. 39, no. 04, pp. 4075–4083, 2012.

[5] G. Susto, A. Schirru, S. Pampuri, S. McLoone, and A. Beghi, "Machine Learning for Predictive Maintenance: A Multiple classifier Approach," *IEEE Trans. Ind. Information*, vol. 11, no. 3, pp. 812–820, 2015.

[6] A. Vallejo-Guevara, J. Nolazco-Flores, R. Morales-Menendez, L. Sucar, and C. Rodríguez, "Tool-Wear Monitoring Based on Continuous Hidden Markov Models," in *Progress in Pattern Recognition, Image Analysis and Applications. Lecture Notes in Computer Science*, no. 3773. Heidelberg: Springer, 2005, pp. 880–890.

[7] A. Vallejo, R. Morales-Menendez, C. Rodriguez, and L. Sucar, "Diagnosis of a Cutting Tool in a Machining Center," in *IEEE Int J Conf on Neural Networks*, 2006, pp. 7097–7101.

[8] A. Vallejo and R. Morales-Menendez, "Cost-Effective Supervisory Control System in Peripheral Milling using HSM," *Annual Reviews in Control*, vol. 34, no. 1, pp. 155–162, 2010.

[9] M. Syafrudin, G. Alfian, N. Fitriyani, and J. Rhee, "Performance Analysis of IoT-based Sensor, Big Data Processing, and Machine Learning Model for Real-Time Monitoring System in Automotive Manufacturing," *Sensors*, vol. 18, no. 9, pp. 2946–2952, 2018.

[10] T. Chen and C. Guestrin, "XGBOOST: A scalable Tree Boosting System," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 785–794.

[11] K. Chen, H. Chen, L. Liu, and S. Chen, "Prediction of Weld Bead Geometry of MAG Welding based on XGBoost Algorithm," *Int. J. Adv. Manuf. Technol*, vol. 101, no. 12, pp. 2283–2295, 2018.

[12] E. Jabbar, P. Besse, J. Loubes, N. Roa, C. Merle, and R. Dettai, "Supervised learning approach for surface-mount device production," in *Proc. Int. Conf. Mach. Learn. Optim. Data Sci*, New York, NY, USA, 2018, pp. 254–263.

[13] Y. Liu and S. Jin, "Application of Bayesian Networks for Diagnostics in the Assembly Process by Considering Small Measurement Data Sets," *Int J Adv Manuf Technol*, vol. 65, pp. 1229–1237, 2011.

[14] E. Wolbrecht, B. Ambrosio, and B. Paach, " Monitoring and Diagnosis of a Multi-stage Manufacturing Process using Bayesian Networks," *AI EDAM*, vol. 14, no. 1, pp. 53–67, 2000.

[15] Y. Liu, S. Jin, Z. Lin, C. Zheng, and Y. K, "Optimal Sensor Placement for Fixture Fault Diagnosis using Bayesian Network," *Assem Autom*, vol. 31, no. 2, pp. 176–181, 2011.

[16] W. Guo and A. Banerjee, "Identification of Key Features using Topological Data Analysis for Accurate Prediction of Manufacturing System Outputs," *J Manuf Syst*, vol. 43, no. 2, pp. 225–234, 1999.

[17] L. Ren, J. Cui, Y. Sun, and X. Cheng, "Multi-Bearing Remaining Useful Life Collaborative Prediction: a Deep Learning Approach," *J Manuf Syst*, vol. 43, no. 2, pp. 248–256, 2017.

[18] R. Sharma, K. Rajagopal, and S. Anand, "A Genetic Algorithm based Approach Forrobust Evaluation of Form Tolerances," *J Manuf Syst*, vol. 19, no. 1, pp. 46–52, 2000.

[19] J. Chen and Y. Liang, "Development of Fuzzy Logic-based Statistical Process Control Chart Pattern Recognition System," *Int. J. Adv. Manuf. Technol.*, vol. 186, no. 1-4, pp. 1011–1026, 2016.

[20] R. Guh, J. Tannock, and C. O'Brien, "IntelliSPC: A hybrid intelligent tool for on-line economical statistical process control," *Expert Syst. Appl.*, vol. 17, no. 3, pp. 195–212, 1999.

[21] C. Lu, Y. Shao, and C. Li, "Recognition of Concurrent Control Chart Patterns by Integrating ICA and SVM," *Appl. Math. Inf. Sci.*, vol. 8, no. 2, pp. 681–687, 2014.

[22] S. Gauri and S. Chakraborty, "Improved Recognition of Control Chart Patterns using Artificial Neural Networks," *Int. J. Adv. Manuf. Technol.*, vol. 36, no. 11-12, pp. 1191–1201, 1993.
[23] C. Liu, K. Chen, S. Jin, and Y. Qu, "Variation Pattern Recognition of the BIW OCMM Online Measurement Data Based on LSTM NN," *IEEE Access*, vol. 7, pp. 69 007–69 014, 2019.
[24] M. Robnik-vSikonja and I. Kononenko, "Theoretical and Empirical Analysis of ReliefF and RReliefF," *Machine Learning*, vol. 53, no. 1-2, pp. 23–69, 2003.
[25] K. Hornik, M. Stinchcombe, and H. White, "Multilayer Feedforward Networks are Universal Approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.
[26] A. F. Amato nd A. López, E. P. na Méndez, P. Vaňhara, A. Hampl, and J. Havel, "Artificial Neural Networks in Medical Diagnosis ," *J of Applied Biomedicine*, vol. 11, no. 2, pp. 47–58, 2013.
[27] P. Valente Klaine, M. Ali Imran, O. Onireti, and R. Demo Souza, "A Survey of Machine Learning Techniques Applied to Self Organizing Cellular Networks," *IEEE Comm Surveys & Tutorials*, p. 1, 2017.
[28] K. G. Sheela and S. N. Deepa, "Review on Methods to Fix Number of Hidden Neurons in Neural Networks," *Mathematical Problems in Eng*, vol. 2013, 2013.
[29] J. Heaton, *Introduction to Neural Networks with Java*. Heaton Research, Inc., 2008.
[30] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, *Neural Network Design*. Martin Hagan, 2014.
[31] C. A. Escobar and R. Morales-Menendez, "Process-monitoring-for-quality—a model selection criterion for shallow neural networks," in *Annual Conference of the PHM Society*, vol. 11, no. 1, 2019.