Cite as: Franciosa P., Sokolov M., Sinha S., Sun T., Ceglarek D., Deep Learning Enhanced Digital Twin for Closed-Loop In-Process Quality Improvement. CIRP Annals-Manufacturing Technology, 69/1, 2020.

CIRP Annals Manufacturing Technology xxxxxxxxxxx

Contents lists available at SciVerse ScienceDirect

CIRP Annals Manufacturing Technology

Journal homepage: www.elsevier.com/locate/cirp

Deep Learning Enhanced Digital Twin for Closed-loop In-Process Quality Improvement

Pasquale Franciosa, Mikhail Sokolov, Sumit Sinha, Tianzhu Sun, Dariusz Ceglarek (1)

Digital Lifecycle Management (DLM), WMG, University of Warwick, Coventry, CV4 7AL, United Kingdom

A digital twin framework is presented for assembly systems with compliant parts fusing sensors with deep learning and CAE simulations. Its underlying concept, 'process capability space,' updates iteratively during evolving tasks of new product introduction with resulting model fidelity able to simulate dimensional and geometric quality of parts and assemblies; isolate root causes of quality defects; and, suggest corrective actions for automatic defects mitigation; thereby, enabling 'closed-loop in-process (CLIP) quality improvement' during assembly system development. Results, using the first fully digitally developed remote laser welding process for Aluminium doors, yielded a right-first-time rate of >96% for door assembly cell development.

Digital Manufacturing System, Assembly, Digital Twin

1. Introduction, motivation and objective

Digital twins are important enablers for current manufacturing systems. Although they strive to represent the physical product and process accurately; currently, the limited accuracy of digital twins of assembly systems prohibit their effective use for simulation of product dimensional and geometric quality [1-2].

For e.g., they are not able to support producing near-zero-defect products and ensure high rate of 'right-first-time' (RFT) [3]. Zero Defect Manufacturing (ZDM) aims to minimize and eliminate product quality defects and process errors. Research in ZDM has led to sensing technology rapidly increasing data generated and captured in-process during installation, ramp-up and full production [4]. However, data gathering and analytics, data-driven monitoring and traditional quality control methods, such as *statistical process control*, though necessary for process monitoring, are understood more as 'open-loops' that lack capability to: (a) isolate root causes of defect(s); (b) identify corrective actions to eliminate isolated root causes; and, (c) develop preventive actions deterring defect occurrence. Moreover, they cannot be currently embedded into digital twin models for use during NPI phases. Hence, it is crucial to (1) enhance current ZDM from data-driven analytics to also include physics-driven CAE models to prevent defect propagation involving multi-stage manufacturing systems. Currently, there is no effective 'closed-loop' solution, which can link defect identification via root cause analysis (RCA) and corrective/preventive actions (CAPA); and, (2) Embed ZDM into the digital twin framework for use in NPI.

This paper presents a digital twin framework with Closed-Loop In-Process (CLIP) quality improvement for assembly systems with compliant parts, which generates CAPA, by fusing sensors with deep learning (DL) and CAE simulations. It fuses process stochastic optimization and variation reduction to enable ZDM with 'RFT'. The CLIP is demonstrated using the first ever fully digitally developed Remote Laser Welding (RLW) process for Al doors.

2. Background

The intensely competitive focus on diversification and responsiveness has led manufacturers to adapt flexible and reconfigurable systems [5]. The ever decreasing time-to-market has forced manufactures to develop new and effective *self-sustainable* systems which must have both the capability to quickly adapt to high product variety and be resilient enough [6] to automatically recover from faulty stages at minimum cost. This essentially translates to producing zero-defect parts faster and cheaper [7]. In this context, ZDM has opened up new opportunities for digital validation during new product and process introduction, along with integrated embedded sensors and data analytics.

The NPI starts with the definition of product and process requirements, along with determining feasible sets of design configurations (*concept stage* (**S1**)). The subsequent *scale-up stage* (**S2**) refines the parameters and engineers the process based on a few (~ few tens) physically representative parts, and entails the detailed design of product and process, and overall system prototyping using pre-production equipment. At **S3**, *production pilot stage*, a full-scale manufacturing system is implemented at the production plant to physically verify and validate design requirements and check process correctability. This is followed by pre-volume production, production launch and full production. This paper focuses on **S1** to **S3** as these are necessary stages to elaborate the concept of closed-loop system implemented as part of the digital twin, and is needed even before full production can start. In fact, any unsolved issues at stages **S1** through **S3** will not only have significant cost increase ('Rule of 10') but also delay product launch due to the incurred engineering changes. For e.g., [8] indicates that most engineering changes for RLW implementation are related to: (i) clamp adjustments - for a door assembly fixture more than hundred changes are made after design release. It takes nearly two weeks to manually set all clamps to satisfy part fit-up requirements; (ii) laser parameters selection and adjustments - it takes up to four weeks to develop a feasible process window to meet all joint requirements; and, (iii) robot path programming - it takes up to four weeks to implement a robot program, which is collision-free and within targeted cycle time.

The NPI process is efficient when the above stages are conducted sequentially with built-in and automatically executed iteration loops. Each iteration loop should take into account the updated information provided by each consecutive stage. Recent research have yielded impressive advancements in digitalisation and virtualisation involving CAE and pertaining to product and process developments. Indeed, a number of studies have integrated current digital tools across the sequential tasks of the NPI. This has led to the framework of Cyber-Physical Production Systems or digital twins [9], with the aim of developing flawless integration between the digital and physical domains, hinging on CAE simulation, real-time data gathering and DL.

3. Closed-Loop In-Process (CLIP) control

A *closed-loop* approach offers capability to accelerate quality maturation and develop corrective strategies for on-the-fly process changes, with integrated data analytics and predictive engineering simulations. Hence, the CLIP must account for:
1. Multi-physics defects: recurring defects propagating through assembly stages trigger complex cause-effect scenarios. As defects are induced by various interactions within the system it is crucial to tackle the dependencies among process parameters and quality faults [7-8]. For e.g., in welding, one clamp may affect mechanical deformation (mechanically-induced fault) while simultaneously impacting heat distribution (thermal-induced fault);
2. High dimensionality and multi-scale data: when dealing with complex assembly systems, the number of parameters may be very large and, due to the multi-physics nature of defects data may spread from 1D signals to 3D scanned surface points. For e.g., in welding process the fundamental mapping between 'macro-scale' strength requirements and 'micro-scale' grain morphology is an essential challenge for selecting process parameters;
3. Stochastic real-time optimisation: Defect propagation can be related to propagation of product and process variations [7] [8] [10] underscoring a need to link real-time optimisation with nested stochastic variation modelling. Existing approaches have mainly focused on modelling product variations without systematic integration with assembly process models; and,
4. Multi-fidelity: multi-physics driven simulations (high-fidelity) are accurate but time consuming. Conversely, data-driven approaches (low-fidelity), based on surrogate/meta-models, are less accurate but allows to infer correlations and dependencies among multiple and coupled parameters in order to reduce the complexity of the design space. Models of different fidelity level are therefore, necessary to leverage the benefits of both data streams. However, establishing the successful trade-off is a challenge.

The CLIP approach is based on the development and integration of three Key Tools (KTs): Multi-dimensional Data Management (*KT1*) [10]; Multi-fidelity Deep Learning Kernel (*KT2*) [11]; and Multi-physics Variation Modeller (*KT3*) [8] [12].

Key Performance Indicators (**KPIs**) are associated to cascading requirements, such as weld quality, dimensional integrity, or cycle time. Key Measurement Characteristics (**KMCs**) measure **KPIs**. The **KMCs** are associated to either real-time in-process data, or knowledge transfer during/after design tasks. Key Control Characteristics (**KCCs**) are process parameters selected to deliver **KPIs**. The j^{th} stage of the NPI is modelled considering {**KPI**(j), **KMC**(j)$^{(t)}$,**KCC**(j)$^{(t)}$, ∀j=1-3}, where 't' refers to the t^{th} iteration loop.

The digital twin is conceptualized to work using different logic flows in different NPI stages. Each of the KTs can be leveraged at different levels. *KT1* can have three levels: No Data (*KT1* at **S**1); Few Production Samples (*KT1* at **S**2) up to 50 samples; Large Production data (*KT1* at **S**3) >100 samples. Similarly, Multi-Fidelity Deep Learning (DL) can be used at different learning complexities: Statistical (*KT2* at **S**1) simple statistical validation using mean and Std Dev; Shallow learning (*KT2* at **S**2) using shallow models such as multiple linear regression or decision trees; Deep learning (*KT2* at **S**3) using deep models such as convolutional neural networks. Simulation can exist at multiple levels depending on complexity: Low Fidelity (*KT3* at **S**3) low complexity models to emulate real system at a macro level; Medium Fidelity (*KT3* at **S**2) medium complexity models to emulate real system considering physics of the process; High Fidelity (*KT3* at **S**3) high complexity models to emulate real-system at a micro level considering multi-physics nature of system including mechanical and thermal effects. Different NPI stages require different levels and combination of various KTs. In the concept stage, *KT1* is at **S**1 given that there is no data hence, *KT2* is leveraged at **S**1 to perform statistical validation of the results obtained by optimisation of different parameters using *KT3* to obtain the system design. At scale-up stage, given there are a few production samples, *KT1* is leveraged at **S2** to give input to *KT3* for it to make an enhanced decision about process variations. This variation data is leveraged by *KT2* to model and, then infer the optimal process design. Finally, in the production pilot stage when there is sufficient amount of data, *KT1* at **S3** is used to train *KT2* at a deep level to perform closed loop control. Given that model performance is not above the required threshold, *KT3* at **S3** (high-fidelity) can be used to generate additional data, then enhance capabilities of *KT2*. All these loops work to optimize various resources of a manufacturing system. While the loops in the early design phases ensure a robust and scalable process design in significantly lesser amount of time, the loops in the production phase ensures CLIP control ensuring higher quality and productivity of the whole system.

The concept of 'uncertainty' is introduced to classify limited knowledge or incomplete data/information during the development stages of the NPI process, caused by the inherent evolution of the process or technological limitations. The first option underlays product development schedules or program management decisions, whereas technological limitations are due to technical barriers to collect data; for e.g., in RLW process, weld porosity and cracks can only be measured using costly destructive tests based on very few samples limiting the applicability for rapid CLIP iterations. Instead, the term 'variation' is adopted in relation to product, process and measurement errors. Intuitively, during NPI, the uncertainty steadily decreases as requirements become clearer; whereas, the source of variation may not decrease due to unforeseen events in early development stages. 'Stochastic uncertainty' is therefore, understood as the quantification of variation with evolving requirements. The stochastic uncertainty at the j^{th} stage is denoted by **ξ**(j).

As demonstrated in [13] an effective approach to quantify stochastic uncertainty is through the concept of stochastic *Process Capability* (*PC*). *PC* gives a quantitative measure of the probability to fulfil input requirements with stochastic variability. As illustrated in Fig. 1, the probability to fulfil the lower (denoted by hat '∨') and upper ('∧') design requirement is formally expressed in Eq. (1), where β is the trade-off confidence value.

$$PC \equiv \mathrm{P}\left(\check{\mathbf{KPI}}(\mathrm{j})^{(t)} \le \left(\mathbf{KPI}(\mathbf{KCC}(\mathrm{j})^{(t)}, \xi(j))^{(t)} \right) \le \hat{\mathbf{KPI}}(\mathrm{j})^{(t)} \right) \ge \beta \tag{1}$$

The *PC* Space is the reassemble of **KCCs** which satisfy Eq. (1). It provides multiple options, each with a different confidence level. This is positive as it provides flexibility when dealing with coupled requirements that involve conflicting control parameters. Trade-off is driven by the desired level of confidence of the solution (i.e., selection of β). The essential part of CLIP is to iteratively calculate instances of the *PC* Space for sequential stages of NPI to achieve:
1. Robustness: it applies when uncertainty is high, for e.g., missing data or input requirements but with fixed level of variability such as pre-determined GD&T. Robustness is typically achieved in the

design stage, where pre-defined assumptions are made on input requirements. The calculated set of **KCCs** is least sensitive to incoming stochastic uncertainty;

2. Scalability: assumptions are now relaxed as sub-set of requirements is settled. Scalability implies the refinement of computed **KCCs** under the new set of assumptions coming from **S2**;

3. Adaptability: it applies when uncertainty is low; for e.g., input requirements are fully defined but with dynamic variability which corresponds to evolving variability over time. Adaptability includes robustness, but not vice-versa. Where dynamic variability is high, robustness, which works with pre-set parameters, does not allow reacting to a range of configurations. This is overcome by adaptability which works with dynamically selected parameters.

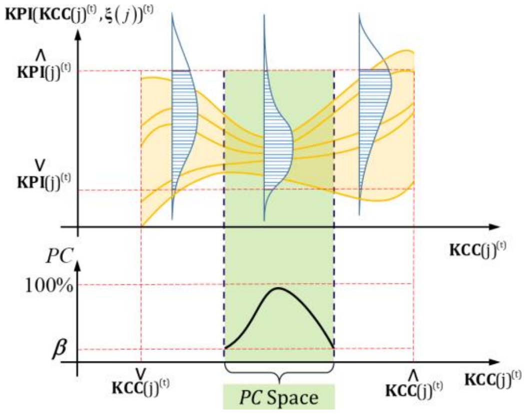


Figure 1. Conceptual definition of Process Capability (*PC*) space

4. Case study: RLW process for automotive Al door assembly

4.1. Background on RLW process

RLW is seen as a key enabling technology in car manufacturing to join lightweight Al structures. Advantages include non-contact and single-sided joining; compared to tactile welding: increased welding speed, reduced flange length (weight), heat affected zone and operational cost [7]. It provides high flexibility as one head can weld multiple products. Still, the uptake of RLW in Al assemblies is limited due to challenges involving process parameters selection related to quality defects (hot cracking, pores) and sensitivity to product and process uncertainties which evolve within NPI. To date, RLW has been implemented in only one automotive OEM AUDI. Hence, to fully exploit its potential, it is crucial to consider the requirements of the whole NPI.

4.2. Concept stage

This includes two major design tasks that interact with each other: (1) selection of welding process parameters, which aims at finding functional mapping between weld quality requirements and welding process parameters; and, (2) conceptual design of assembly fixture optimising clamp layout. The interaction between (1) and (2) is driven by the variability of part-to-part gap. For instance, in automotive body-in-white sheet metal assembly, form tolerances on stamped parts may raise up to ±0.5 mm, which leads to part-to-part gaps of up to 1 mm. The absence of filling material and mechanical guns which tend to close the fit-up gaps, are a challenge to compensate manufacturing tolerances [7]. Part-to-part gap can be compensated either by optimising the number and position of clamps (clamp layout optimisation), or selecting the welding parameters, such as laser power and welding speed, such that they are least sensitive to gap variations. Trade-off decision between design tasks (1) and (2) is reached by considering downstream functional requirements such as dimensional quality and strength of the sub-assembly: over-constraining the parts helps to close the gaps but leads to unwanted residual stresses and spring-backs ('dimensional quality' requirement); higher gaps can be controlled with optimised welding process parameters; however, as a result the weld strength may be hindered by the reduced cross-section length ('strength' requirement).

Stochastic uncertainty **ξ**(1) is as follows: (a) at concept stage, production requirements are partially defined. Hence, target cycle time, which drives welding speed, and manufacturing tolerances (i.e., part-to-part gaps) rely on 'best-known' assumptions; (b) stamped parts are unavailable, hence, predictive models can be built only with historical data or GD&T tolerances; and, (c) only surrogate parts in the form of flat coupons can be used to release the preliminary set of feasible welding parameters. Input parameters, **KPI**(1), are: material alloy specifications and thickness, expected joint geometry (i.e., fillet or overlap weld), estimated welding speed, and available maximum laser power. **KMC**$(1)^{(t)}$ are iteratively generated with physical experimentation (i.e., coupon testing), which is driven by the CLIP kernel that helps selecting the next set of control parameters, **KCC**$(1)^{(t)}$. Control parameters are associated to welding parameters (i.e., laser power, scanning amplitude) and number/position of clamps.

Determinations of *PC* Space are presented for both design tasks (1) and (2). Note is made that at this stage of the NPI, the 3D CAD model of the assembly fixture is unavailable. Thus, the clamp geometry is approximated with cylindrical or prismatic shapes, corresponding to the footprint of the clamps on the parts. Detailed geometry can then be obtained in the subsequent scale-up stage. Fig. 2(a) shows the *PC* chart related to design task (1), with confidence level of 80%. Blank areas in the *PC* chart correspond to unfeasible sets of process parameters. Higher values correspond to high likelihood of meeting requirements. For e.g., the sample in Fig. 2(a)(1) shows a weld cross-section indicating high robustness against gap variation (*PC*>80%). Notice that the *PC* space is very narrow and samples in Figs. 2(a) (2) & (3) show considerable reduction in weld quality resulting in 'no weld' fault (*PC*<80%).

Similarly, Fig. 2(b) shows the *PC* chart for design task (2). The study evaluated robustness of clamp positions (all clamps have been moved parallel to welding flange for the same amount, in the range [-5,5]mm) and number of clamps against form errors of parts being welded: door inner panel and upper window reinforcement. Results indicate a wide *PC* space. Fig. 2(c) shows the corresponding clamp layout, where the contour plot is the part-to-part gap. The highlighted regions (1)-(4) correspond to potential failures, as confirmed quantitatively (Fig. 2(d)).

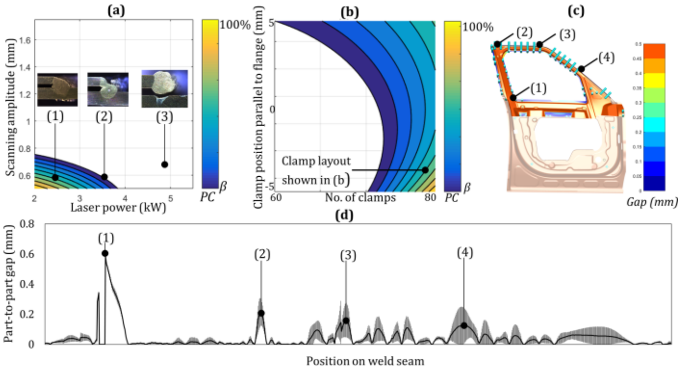
4.3. Scale-up stage

Scale-up refines the process parameters computed at concept stage by verifying and validating it using physical components and pre-production equipment. *Scalability* entails transition from surrogate to representative parts; and, development of assembly cell using pre-production equipment. Input parameters, **KPI**(2), correspond to pre-computed process parameters in **S1** refined by using multi-scale data, **KMC**$(2)^{(t)}$, such as scanned 3D point clouds, surface temperature, weld grain morphology, mechanical strength.

Stochastic uncertainty, **ξ**(2), is now associated to the limited dimensionality of the available dataset. For instance, during scale-up the first set of prototype parts becomes available, but with limited sample size – only few tens. Also, as they come from prototype press-shop, GD&T specifications are not met fully.

Fig. 3 shows the effect of transition from concept to scale-up. Applying the set of **KPI**(1) directly in **S2** without any further refinement leads to severe weld cracks (Fig. 3(b)). This is due to the change in flange design, which generates new solidification and heat dissipation mechanisms, not predicted during the **S1** stage.

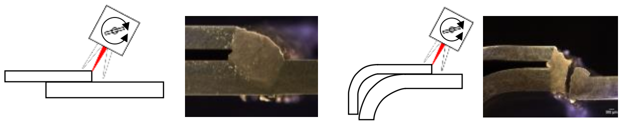
A multi-scale data approach is implemented at this stage to rapidly scale process parameters. Since the properties of the weld are closely dependent upon its microstructure, the cooling and solidification rates of the weld strongly influences the final macro-scale quality indicators. For e.g., Fig. 4 illustrates the relation between cooling rate, measured with thermal camera, and grain morphology, using EBSD microscopy. Data have been segmented and clustered using K-means clustering Results show that the highest cooling rate is achieved for the **S1** flange design. Consequently, when moving to the **S2** design, grains are on average bigger due to lower heat dissipation. As grain size is directly related to crack sensitivity, when refining the process parameters, the net input energy to the material is reduced by 15% (Fig. 4).



(a) *Welding process parameters* selection with part-to -part gap assumed N(0.5mm, 0.1mm). Conf β=80%
(b) *Fixture layout* design for part form error assumed N(0mm, 0.05mm). Conf. β=80% w/ max gap 0.5mm
(c) Clamp Layout of fixture design for (b)
(d) Calculated part-to-part gap variation for (c)

Figure 2. *Process Capability (PC)* Space at Concept Stage for (a) welding process parameters selection, and (b) fixture layout design

The detailed assembly fixture configuration has been achieved with CLIP iterations which involve simulation of robot visibility and accessibility, along with cycle time optimisation, and continuous refinements of the *PC* Space (presented in Sect 4.2) with 3D scanned points of prototype parts. These automatically executed iterations helped to improve the confidence towards optimal design. Fig. 5 shows the transition steps from conceptual design to pre-production fixture.



(a) *Concept* stage: Flange for coupon testing (b) *Scale-up* stage: Final flange design

Figure 3. Transition from (a) surrogate parts to (b) representative parts

4.4. Production pilot stage

Fig. 6 shows the RFT rate achieved during the production trials. The trials ran for eight days in the assembly plants producing nearly 3000 sub-assemblies. The RFT rate has been calculated considering the amount of welding length, which is defect-free (i.e., no cracks, no seam discontinuity, and no lack of gap bridging). Results show that after commissioning, the achieved RFT rate is more than 96%. This is a very positive result since based on industry quality standards 15% extra welding length is added to the sub-assembly 'just-in-case' to compensate those joint defects, which cannot be detected and eliminated. This result impacts both cost reduction and also increased throughout (with the same cycle time), and reduction of manual repair actions.

5. Conclusions and final remarks

Current ZDM aims to reduce or eliminate product and process failures to achieve 'right-first-time'. Best practices are based on 'open-loops' which attempt to link process monitoring to process adjustment, via trial-and-error. This practice is well established for low volume production systems such as aerospace, is unviable for large volume productions, as in automotive industry.

Hence, a paradigm shift is necessary to address the shortcomings of conventional approaches. The presented CLIP approach fuses in-process data, data analytics and physics-driven simulation to diagnose defects and correcting and preventing their occurrence. It is based on the underlying concept of stochastic process capability space, which is iteratively updated during the evolving NPI stages. Results underscore a need to develop a system which is *robust*, *scalable* and *adaptable* to cope with dynamic changes, unforeseen at early stages of NPI.

The benefits of the CLIP approach are: (1) Faster selection of process parameters, which currently takes several months of physical experimentation. The presented model can help to reduce overall number of physical experiments; (2) Capability to automatically adjust process parameters by leveraging stochastic uncertainty during the NPI stages; and, (3) Real-time closed-loop gap bridging control with adaptive selection of new set of process parameters. Results of the production pilot study yielded an impressive right-first-time rate of >96%.

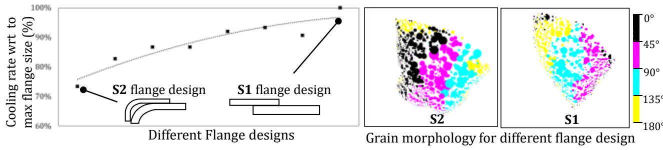


Figure 4. Scale-up of process parameters wrt. flange design conditions

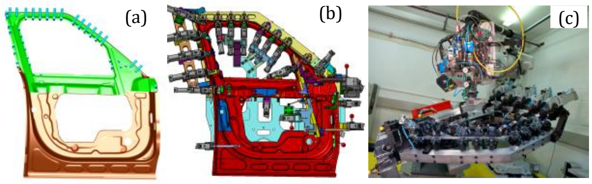


Figure 5. Fixture Scalability: (a) concept (b) detailed; (c) pre-production

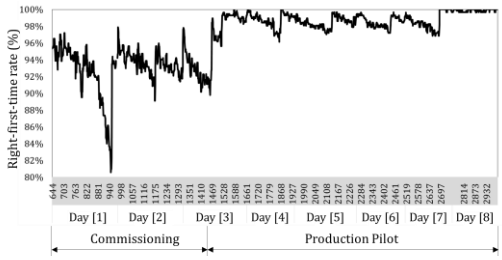


Figure 6. Right-first-time rate achieved during production pilot study

6. Acknowledgments

This study was partially supported by the WMG Centre HVM Catapult, APC UK: Chamaeleon, and the EPSRC EP/K019368/1.

7. References

[1] Söderberg, R, et al., 2017, Toward a Digital Twin for real-time geometry assurance in individualized production, Annals of the CIRP, 66/1:137-140.
[2] Schleich, B., Anwer, N., Mathieu, L., Wartzack, S., 2017, Shaping the Digital Twin for Design and Production Engineering, Annals of the CIRP. 66/1:141–144.
[3] Lee, J., Lapira, E., Yang, S., Kao, A., 2013, Predictive Manufacturing System – Trends of Next-Generation Production Systems, *11th IFAC Workshop on Intelligent Manufacturing Systems*, 46/7:150–156.
[4] Wang, K., 2013, Towards Zero Defect Manufacturing (ZDM)—a Data Mining Approach, *Adv. Manuf.*, 1:62–74.
[5] Koren, Y., Jovane, F., Moriwaki, T., Pritschow, G., Ulsoy, G., Van Brussel, H., 1999, Reconfigurable Manufacturing Systems," Annals of the CIRP, 48/2:527-540.
[6] Ceglarek, D., et al., 2004, Time-Based Competition in Multistage Manufacturing: Stream-of-Variation Analysis Methodology, *Int. J. Flex. Manuf. Syst.*, 16:11–44.
[7] Ceglarek, D., *et al.*, 2015, "Rapid Deployment of Remote Laser Welding Processes in Automotive Assembly Systems," Annals of the CIRP, 64/1:389–394.

[8] Franciosa, P., Ceglarek, D., 2015, "Hierarchical Synthesis of Multi-level Design Parameters in Assembly System," Annals of the CIRP, 64/1:149–152.

[9] Uhlemann, T., Lehmann, C., Steinhilper, R., 2017, The Digital Twin : Realizing the Cyber-Physical Production System for Industry 4.0, Procedia CIRP, 61:335–340.

[10] Babu, M., Franciosa, P., Ceglarek, D., 2019, Spatio-Temporal Adaptive Sampling for effective coverage measurement planning during quality inspection of free form surfaces using robotic 3D optical scanner, J. Manuf. Sys, 53:93-108.

[11] Sinha, S., Franciosa, P., Ceglarek, D., 2020, Object Shape Error Response using Bayesian 3D Convolutional Neural Networks for Root Cause Analysis of Assembly Systems, submitted to IEEE Trans on Industrial Informatics.

[12] Franciosa, P., Palit, A., Gerbino, S., Ceglarek, D., 2019, A novel hybrid shell element formulation (QUAD+ and TRIA+): A benchmarking and comparative study, Finite Elem. Anal. Des., 166:103319.

[13] Franciosa, P., Gerbino, S., Ceglarek, D., 2016, Fixture Capability Optimisation for Early-stage Design of Assembly System with Compliant Parts Using Nested Polynomial Chaos Expansion, Procedia CIRP, 48th CIRP CMS, 41: 87-92.