Quality Monitoring and Fault Detection in an Automated Manufacturing System – a Soft Computing Approach

Mika Ruusunen and Marko Paavola

Report A No 19, May 2002
QUALITY MONITORING AND FAULT DETECTION IN AN AUTOMATED MANUFACTURING SYSTEM - A SOFT COMPUTING APPROACH

Mika Ruusunen and Marko Paavola

Abstract: Quality monitoring and fault detection are essential parts in automated electronics manufacturing systems. Information about process conditions enables operations to improve quality and increase throughput. This report presents a general quality monitoring framework and method for a manufacturing system.

Proposed monitoring approach is an integration of model-based methods with systematically collected expert knowledge and data. A model bank is constructed to reproduce behaviour of the normal and fault states. The data driven normal condition model contains linguistic equation - non-linear scaling method for model variables, and a recursive gradient algorithm. Fuzzy reasoning and basic statistical methods are combined to identify changes in normal model residuals. Fault models are fuzzy rules for detecting abnormalities in selected time series signal. Analysed model outputs are then applied to monitoring task.

Principles of the monitoring method are briefly discussed and demonstrated with a simulation example. Modelling results indicate that the proposed method can handle noise in simulation data. Generalisation ability of the normal model was also notified. Based on simulations, presented monitoring approach was verified to have potential features to be implemented as a real time application.

The work is part of the INTELE - Intelligent Methods in Electronics Manufacturing research project, financially supported by Tekes, Filtronic LK Oy, JOT Automation Oy and PKC Group Oyj. INTELE project belongs to Tekes technology programme ÄLY - Intelligent Automation Systems 2001-2004.

Keywords: Electronics manufacturing, condition monitoring, quality control, fuzzy, neural.
INTRODUCTION

PROCESS MONITORING - DEFINITIONS

PROPOSED GENERAL MONITORING FRAMEWORK

CASE: SCREW INSERTION SYSTEM

4.1 Process description

4.2 Literature review: Monitoring methods for screw insertion

4.3 Monitoring approach for screw insertion process: overview

4.4 Design of experiments

SIMULATION EXAMPLE

5.1 Simulation data and structure of the monitoring software

5.2 Quality monitoring: Normal condition model

5.2.1 Model structure

5.2.2 Training and validation

5.2.3 Monitoring and prediction of process quality using fuzzy thresholds

5.3 Fault detection and isolation: Fault models

5.4 Comparative study: principal components and neural networks

CONCLUSIONS

REFERENCES

APPENDIX
1 INTRODUCTION

A modern end-of-line electronics manufacturing environment is a combination of individual cells, designated to complex assembly and material processing tasks. The success of such a system is greatly depended on fault-free and stable operation of every unit. Deviations from normal situation decrease productivity. Continuous monitoring is therefore required to maintain the functionality and quality of the production.

Reliable detection and isolation of faults is an important part in a successive maximisation of productivity. It provides the needed information for fault diagnostic and error recovery operations. Connected with continuous monitoring, estimating the evolution of the failures becomes possible and predictive maintenance policy can be achieved. Fault detection and quality control of manufacturing processes has traditionally based on limit value checking of measured signals. Flexible manufacturing environment makes it difficult to use constant limits for alarms. Interpretation of monitoring results is also complicated due to uncertainties in collected information. Moreover, thresholds that define the normal conditions are often dependent on the operating point of the system. Detection and isolation of small, simultaneous faults is therefore difficult with signal-based limits.

Model based fault detection methods have been used in production industry to overcome difficulties that arise with limit checking. These methods have many advantages, for example a higher performance - smaller faults can be detected and different faults can be isolated. Disadvantages are that an accurate model of the process is necessary for efficient operation /23/. In recent years, many soft computing methods have been introduced for fault detection. Neural networks and fuzzy logic are typical examples of artificial intelligence methods, generally model-based reasoning /41/. Fuzzy set theory, for example, is an effective tool when dealing with uncertainty in systems /24/. Fuzzy rules are easy to interpret and causal relations of the system can be presented in the form of natural language. Neural networks are data-driven methods with the ability to generalise on the basis of collected information /13/. In addition, linguistic equations framework has been introduced and applied for several process control and monitoring tasks (/19/, /20/, /21/).

In this work, various different monitoring strategies are studied and combined in order to achieve a general fault detection and quality monitoring framework. Possibilities to improve early detection of abnormal process conditions in manufacturing cells are investigated. To reach the targets, intelligent hybrid methods are considered and applied together with traditional statistical control approaches, including expert knowledge. In the proposed framework, significant faults of the selected application are first explored via systematic analysis technique. Models for normal and fault situations are then constructed on the basis of analysis results, and information collected from designed experiments. Finally, outputs of the models are interpreted in order to identify faults and production quality. Functioning of the method in different process environments is ensured using selected analysis techniques and parameter adaptation. Principle of the developed monitoring system is demonstrated with simulated data.
2 PROCESS MONITORING - DEFINITIONS

Monitoring is a continuous real-time task of determining the conditions in a physical system. It consists of recording information, recognising changes and detecting abnormalities in the system's behaviour /14/.

Fault is a deviation of at least one characteristic property or parameter from an acceptable condition. It is a state, which may lead to a malfunction or failure of the system /16/. Typical fault-classes in function of time are presented in Figure 1. Abrupt fault appears as a quick increase in observed symptoms, for example a sudden breakdown of component. Incipient fault gradually develops as a function of time; a typical example is degradation of a tool wear. Intermittent fault repeatedly occurs and disappears /23/.

**Figure 1.** Time-dependency of faults /16/.

Detection of faults is based on symptoms that are changes of observable quantities from their normal behaviour. Symptoms can be generated from analytic and heuristic information. Analytic symptoms consist of characteristic values of the system. Measurement data processing has to be performed in order to construct these values. Various methods are used, namely /16/:

- limit value checking of directly measured signals. The characteristics values are exceeded signal tolerances,
- signal analysis of measurements with signal models like spectral analysis or frequencies, variances, amplitudes and model parameters,
- process analysis using mathematical process models, parameter estimation, state estimation and parity equation methods. In this case characteristic values are model parameters, state variables and residuals of the models.
The resulting changes in characteristic values can be considered as analytic symptoms. Heuristic symptoms, on the other hand, are based on qualitative information obtained from human operators. Further sources for producing heuristic symptoms are process history, maintenance reports, statistical data, lifetime and load measures /16/.

Fault detection includes recognition of an unacceptable behaviour in a system. Abnormal conditions can be detected with analysing existing symptoms that are observable notifications of fault situation. Fault isolation, that is the determination of the type and location of the fault, is performed after detection. Strategies for monitoring process abnormalities are called Fault Detection and Isolation (FDI) methods. /16/

Quality monitoring of production systems includes observation of the product quality, process quality and functioning of machines. Also the reporting can be considered as a monitoring method. Information about process conditions and quality data enables the analysis and implementation of process and quality control mechanisms /5/. In the following chapters a general framework for monitoring of automated electronics manufacturing systems is presented. Principles for proposed solution are derived from literature and case examples by combining different methods.
3 PROPOSED GENERAL MONITORING FRAMEWORK

The general framework proposal for quality monitoring and fault detection originates from the company’s need to continuously improve the process and product quality, including throughput of its production units. Observation of these quality measures is in major role when new process and quality control mechanisms are to be implemented.

Unlike in conventional monitoring strategies, where methods are case specific, the framework should be as general and flexible as possible. Monitoring principles should also be able to standardise in order to achieve a portable structure. Based on the demands, the general monitoring framework is outlined and presented in Figure 2.

![Diagram of Proposed General Quality Monitoring and FDI Framework](image)

**Figure 2.** Proposed general quality monitoring and FDI (Fault Detection and Isolation) framework for manufacturing systems.

The first task in the proposed approach is to find significant symptoms that are robust against noise, disturbances, uncertainties and set point changes of the process. Failure Modes and Effects Analysis (FMEA) /35/ can be applied for this stage together with expert knowledge (Figure 2). Both process and machine FMEA are in occasional use in the company at this moment. The use of FMEA can help to form a general view to a problem under consideration. Systematic exploration of common failure modes, their effects and causes makes the selection procedure of robust symptoms more reliable. Collected information can also be utilised to define the main focus areas and generic knowledge base of fault detection part.

After symptom selection, the needed measurements have to be defined. Since one of the goals is to keep the number of sensors low for economical reasons, special attention should be drawn on this stage. Measurements are chosen on the basis of previously specified symptoms, which may increase the number of sensors. Monitoring in flexible assembly operations can be also difficult using only sensor related information. To reduce sensors and still maintain the high process information level, a model-based approach for monitoring task is suggested. In model based monitoring, the specific
behaviour of the system is simulated with a model. Deviations in the model accuracy, parameters, outputs or state variables are then detected to decide whether process conditions are changed /14/. The following general advantages of model based diagnosis systems are presented in literature compared to monitoring with sensors and limit checking only /34/: 

- Smaller and also more different types of faults can be detected and the detection time is shorter.
- Detection can be performed over a large operating range.
- Model based detection is a passive procedure, without disturbing the operation of the process itself.
- Increases possibilities to perform fault isolation.
- Disturbances can be compensated for, which implies that high diagnosis performance can be obtained in spite of disturbances.
- Cost and weight of the monitoring system can be reduced.

The disadvantage of model-based fault detection is the need for an accurate model, and usually a more complex design procedure. Deeper insight and understanding of the process is also required. When dealing with highly changing and noisy process conditions with limited amount of data available, conventional modelling methods often face great difficulties. Computational intelligence methods such as neural networks and fuzzy systems offer a way to cope with these problems, and are increasingly probed in production industry nowadays: /1/, /12/, /15/, /19/, /21/, /24/, /27/, /41/. An adaptive model structure can cope with the demand of the flexible monitoring system, and may be transferable to other process environments with a small amount of additional resources. One of the desired aspects of intelligent methods is their reported capability to generalise information from data. This calls the need of representative data for model identification.

Design of experiments (DOE) /2/ is a methodology to plan systematic tests for collecting statistically significant data from the process. Designed experiments can help to propagate knowledge about behaviour of the selected symptoms in different situations. When included in the presented monitoring framework (Figure 2), it guarantees the informative nature of the data for model identification. If DOE is first performed in processes where the monitoring system is to be transferred, goodness of identification data is ensured every time. Also the optimisation of processes becomes possible, since targets of the DOE can be set to explore optimal control parameters of the system as well.

Model bank of different process conditions is proposed. It is a definition for a group of separate models that produce the parallel output. The modelled system can be divided this way into simpler parts, making the modelling task easier and more accurate /42/. The concept of model bank could suit to the monitoring task, since differences between observed normal and abnormal situations may interfere the performance of a single model.

In suggested monitoring framework (Figure 2), both normal condition model and multiple fault models are presented. Model of normal behaviour is intended for quality
monitoring. It is trained using DOE-data, resulting a model that may be general but sensitive to abnormal changes. Connected with other quality parameters (defined at symptom selection stage), observing of the modelling error is then applied for process and product quality monitoring. It may also be appropriate for prediction of machine parts wearing. Adaptation of the model to new situations can be assured with on-line parameter tuning. The detailed scheme of the suggested model based monitoring approach is shown in Figure 3.

Figure 3. Detailed structure of the suggested model-based monitoring approach. SPC stands for statistical methods; FDI is fault detection and isolation stage.

The information available in the results of failure mode and effects analysis includes partially vague and inexact, often conflicting expert knowledge. This kind of information is difficult to utilise in traditional expert systems, but favours fuzzy logic approach in the form of rule base /37/. Fault model are therefore suggested to be constructed using fuzzified expert knowledge where possible. Rule based monitoring of faults is also acceptable in situations where failure data is missing. If several characteristic values are available, different faults can be isolated using classification methods or known fault-symptom causalities. This allows mapping of specific symptom patterns to individual faults.

Outputs of the model bank have to be analysed to make conclusions of the system conditions. One of the main advantages of the model based monitoring approach is that if model explains a process accurately, the modelling error value has a constant scale and variation. Model can be interpreted as a filter that transforms the correlated input data to uncorrelated data, and statistical may be applied to the latter in order to detect changes or the presence of out-of-control signals /38/. Using fuzzified continuous values of the observed symptoms, the instant changes in quality can be monitored. Next, detailed aspects of the presented monitoring framework are discussed with the help of case study.
4 CASE: SCREW INSERTION SYSTEM

4.1 Process description

Screw insertion is a typical assembly task at manufacturing line. The main target of this process stage is the joining of two parts together with screws, using desired force. The insertion process can be described by dividing it into four stages (Figure 4).

The first stage lasts from the first contact of the screw with the hole until the cutting portion has been completely inserted into the hole. During the second stage screw breaks through the material and the required torque decreases. The third stage, screw advance, lasts from the screw breakthrough until the screw cap makes first contact with the plate.

On stages one, two and three the main forces involved are shear forces due to thread forming, and friction forces due to screw movement. Once the screw breaks through, the shear or cutting forces become negligible and friction forces dominate. Screw tightening (stage 4) starts after the screw cap has made contact with the near plate. It finishes when the predetermined clamping force is applied to the joint. At the last stage the tightening forces are involved. /25/

The increase in the strength of the material results an increase in the torque necessary to overcome the generated forces. The geometrical properties of the screw and the plates also determine the torque during the insertion. Three parameters have a high impact on the shape of the torque signal: material, screw size and hole diameter /25/. In the next section, several previously proposed monitoring methods for screw insertion process are discussed.
4.2 Literature review: Monitoring methods for screw insertion

Monitoring of the screw insertion process is currently based most widely on the Teach Method. The strategy is to compare real time torque versus rotational angle signals with the known signature signals /40/. The correct signature signals are taught prior to production. This could involve lengthy set up times and lead to lack of flexibility /39/.

Application of artificial neural networks for monitoring of self-tapping screw insertion is presented in /25/. A radial basis function network is employed to distinguish between successful and failed insertions. It is concluded in the paper that torque signal contains descriptive information for monitoring the process. Time series of the torque signal from individual screw insertions were used as inputs for the neural network. The method was first trained and tested with simulation data based on analytical models of the screw insertion. Data sets using various different process parameters were generated. Network was then trained with a real data collected from experiments, where a hand held screw drive was used. Again, process conditions were varied including hole diameter, plate materials and screw sizes. Training data contained both successful and unsuccessful insertions. A 100% accurate classification between successful and failed insertions without fault isolation was achieved. Preliminary tests indicated that when the network was introduced with unseen data, it was capable to generalise on the basis of training data.

Thread forming can be considered similar to screw insertion, what becomes to dynamic behaviour of the torque signals. In /6/, three methods are presented for monitoring the quality of the thread forming process: Neural networks combined with genetic algorithms, fuzzy logic and fuzzy clustering. According to reference, tool wear is a major symptom in a quality prediction of the threads. Using different features of torque signals as inputs for discussed methods, quality index of the process was generated. Features for classification with neural networks included for example multiple slopes, maximum value, inflection points and area above the zero axis of the torque signal. The use of ten features yielded best results. Achieved recognition rates were around 82.3 %. After the network structure was optimised with genetic algorithms, recognition rate of a tool wear increased to 90.9 %.

The fuzzy logic approach for quality monitoring presented in /6/ consists of a rule base, generated using expert knowledge. The symptoms selected to inputs for final rule base were the maximum of the torsion course, its standard deviation on the basis of the last ten measurements, the first zero of the torsion course, and the number of threads formed. Output variable was defined as degree of wear, which was considered as a quality measure for the produced threads. To reduce number of the rules, the inputs were connected into a two partial rule so that the rule base could be visualised in two dimensions. Connected input variables were then plotted versus each other. This way it was possible to monitor the relation of two variables as a function of the operation time. Tests with real data showed that the fuzzy logic based method for quality monitoring can be realised by heuristically constructing a fuzzy rule base for this case.
Fuzzy c-means clustering for thread forming monitoring was tested with the same data as previously discussed methods in /6/. The maximum torsion course, standard deviation of the signal and first zero of the torsion course were chosen as input variables. Two clusters were labelled, namely “low wear” and “high wear”. Performed tests with data showed that membership values of the “high wear” increased with the number of forming events. The results indicated that fuzzy clustering method can be considered to monitor process, for example to stop the process before the poor quality threads are formed. It was concluded that the tested clustering method is very useful if only a rough knowledge of the system’s properties is available. This is true since the method doesn’t need explicit expert’s knowledge. The adaptation time is shorter in this case also, compared to above mentioned neural networks and fuzzy logic system.

Paper /32/ presents initial simulation results of an intelligent screw insertion control system, focused on sensitive assembly operations. The developed system supervises equipments by giving continuously instructions to control the screw insertion process. The system includes a fuzzy feedback controller containing specifications about successful insertions in the form of a rule base. Target of the control approach is to maintain an acceptable level of joint torque at the corresponding angular position or depth of the screw. The rule base is developed based on the understanding of how the process works. The fuzzy system was tuned heuristically and tested with simulated data. Based on simulation results, the authors verified the usefulness of the fuzzy logic controller in adjusting the driver motion to guarantee the acceptable joint torque. It was also notified that the same control scheme is capable of early error detection, for example in jamming or slippage situations. /32/

Weightless neural networks (WNN) based monitoring strategy for automated self-tapping screw insertions is proposed in /39/. The network is first trained and tested using computer simulations. After that model is tested with experimental data. Trained model is used for classification task between two classes: normal and unsuccessful screw insertions. WNNs are conceptually different from conventional neural nets. The training data of WNNs is stored in a memory, whereas with conventional networks the data is used to adjust the weights. Network consists of Single General Neurons (SGN), comparable of conventional neural networks. SGN includes address decoder, memory registers, data-input and data-output registers. The neuron has two working modes: training and recalling. During training every SGN receives, processes and saves information for future reference. In recall, training information stored in SGN is applied to reach output. When the input vector is received, the corresponding memory register address is located, and data stored at this address during training is given as output. Detailed structure of WNN is given in /39/. Test results showed that the main advantage of using WNNs was their ability to generalise and cope with unseen situations. The network was also relatively easy to train and implement. The authors indicated the promising potential of WNNs for the monitoring of screw insertions.
4.3 Monitoring approach for screw insertion process: overview

The proposed monitoring approach of the automated screw insertion unit follows the guidelines of the presented general framework (Figure 2). Significant faults are first identified with FMEA. Effects of faults in screw insertion are mostly related to defects in product. For this reason the insertion stage is considered for monitoring purposes. Principle of the designed monitoring approach is shown in Figure 5. Measured and modelled signals are used as input to model bank that includes both normal condition and fault models. Faults to be monitored are chosen according to failure analysis. Quality monitoring and fault detection are carried out by jointly interpreting model outputs.

![Figure 5. Principle of the monitoring approach for screw insertion system.](image)

Symptom selection follows the fault identification. Deviations of the monitored signal can be regarded as a robust symptom against changes in the set point and normal process conditions /39/. According to /44/, monitoring of torque and rotational angle signals is the standard for high performance screw insertion. Information about these variables is commonly used for statistical analysis of failure rates /33/.

Also other symptoms have been identified for indicators of any quality problems. According to literature and expert knowledge, deviations in the following quantities can be considered as symptoms for quality monitoring and fault detection:

- Model residuals
- Behaviour of the torque signal
- Set value of the screw advancement
- Set point torque average
- A real-time revolution of the screw drive
- Fault frequency
- Set point torque standard deviation
- Cycle time (product specific) /31/
- Dimensions and space between joined parts (product specific)
- Product and material information (product specific)
By using installed measurements of a screw drive, a certain signal could be modelled (signal model in Figure 5). Thus the model would predict the value of quality parameter instead of a costly sensor. When constructing a normal condition model (Figure 3) for prediction of successful insertions and then comparing modelled output to measured or modelled signal value, deviations from normal behaviour could be observed (section 3). In this case, modelling error acts as a symptom. Modelling methodology also offers a way to monitor the quality of the process/products if the re-trainable model is instructed to imitate the high quality screw insertions.

Following the monitoring framework to the measurement selection stage (Figure 2), measurement methods for specified symptoms may include:

- Modelled signals, based on measurements (Figure 5)
- Acoustic emission measurements /7/
- Vibration measurements /36/
- Acceleration sensors
- Force sensors
- Vision monitoring systems (2D, 3D) /45/

Combinations of these measurements could be used as inputs for normal condition model, or in some cases independently. Also achievements in development of small micro electromechanical systems (MEMS) are bringing new possibilities to integrated measurement implementations /36/.

In the future research, one of key areas is the predictive monitoring of tool wearing and compressed air screw drive degradation. Following methods have been investigated in advance, namely:

- Time series trend of the normal condition model residuals connected with SPC.
- Monitoring the number of produced products and tool lifetime relation.
- Estimating lifetime of parts by fitting acceleration models to degradation data/10/.

Early detection of degradation in machine parameters could be realised by utilising traditional SPC tools to detect changes in normal condition model error. Accelerated life tests /9/ can be performed to collect data for modelling of machine part degradation. For successful monitoring of the tool wear, there must be indications of incipient faults (section 2) in the collected data. Up or down drifts in data are common symptoms for degradation.

For identification of the models, design of experiments has to be planned. Next section presents an approach to systematic data collection. A non-standard design is applied to this case since there are multiple and biased targets for the experiments making the design procedure difficult with traditional design procedures.
4.4 Design of experiments

Based on monitoring framework presented in section 3, experimental design for screw insertion process was planned. In this case the systematic design of experiments was required to achieve three targets:

1. Representative data for modelling.
2. Knowledge about effects of single variables to process behaviour.
3. Optimisation of the process parameters: finding optimal and robust settings of selected control parameters for assuring high quality and throughput.

Selection of variables to be used in experimental design was made in cooperation with process experts. Only the most significant variables, called factors, were chosen. To achieve information from various operation conditions, several different levels were defined for each factor. Three levels were specified for all but one of the factors that was planned to have two levels. Multilevel design makes it possible to identify nonlinearities between process response and each of the factors /30/.

Unfortunately, multiple level standard designs usually require a high number of runs. At the same time, cost and efforts may restrict the amount of experiments. Assuming that a non-linear model structure is needed, standard designs may fail. To realise a mixed two and three level experimental set up with minimum runs, an optimal design procedure is chosen. Optimal design is a general definition for a group of algorithms designated to find an optimal design matrix according to some criteria under specific constrains. Orthogonality of the matrix is often used as a criterion to determine if the solution is good enough. If design matrix is orthogonal, it allows interpretation of the effect estimation of certain factor independently of the other factors influence. Thus the main effect of each individual factor can be observed /29/.

For this problem, optimal experimental design is constructed utilising a columnwise-pairwise (CP) algorithm /28/. It has proven to be efficient and flexible method in mixed-level designs /29/. CP algorithm optimises a design matrix by manipulating its column pairs so that a combination, which gives the highest criterion, is finally chosen. In this case, a normalised D-criterion /43/ is used to evaluate the design. Selected factors with their levels are presented in Table 1.

Table 1. Factor levels.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Low level</th>
<th>Medium level</th>
<th>High level</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>Yes</td>
<td>No</td>
<td>-</td>
</tr>
<tr>
<td>x2</td>
<td>Soft</td>
<td>Medium</td>
<td>Solid</td>
</tr>
<tr>
<td>x3</td>
<td>0.9</td>
<td>1</td>
<td>1.01</td>
</tr>
<tr>
<td>x4</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
</tbody>
</table>
An optimal mixed-level design constructed by CP algorithms is shown in Table 2. Design matrix was made using public software, programmed by William Li, the developer of the CP-algorithms. Numbers in the design matrix present levels 1, 2 and 3 of each factor. The experimental sequence is randomised to avoid any biases during experiments.

Table 2. Completely randomised Mixed-level Design, normalised D-criterion /29/: 0.998.

<table>
<thead>
<tr>
<th>Experiment number</th>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>x4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Normalised D-criterion of the design is 0.998, meaning that the orthogonality criteria is almost fulfilled. The design can be replicated by performing each combination of factor levels more than once. This will allow variability estimates of measurements. Experiments should be carried out following the design matrix. After the data collection it should be pre-processed in order to get faultless data sets.

The collected data from presented experimental design can be first analysed following the Taguchi methodology. By calculating averages and signal to noise (S/N) ratio of the factors with respect to selected response variable, interactions can be analysed. For S/N ratio, target specifications have to be defined. If optimisation goals during experiments are going to be fulfilled, confirmation runs are needed to test the goodness of control parameters (combination) set values.

The collected data is ready to be used for model identification right after pre-processing. In the next section, basic ideas of the normal condition and fault models are presented and tested with simulation example. Developed models are used for quality monitoring and fault detection together with fault isolation of the screw insertion process.
5 SIMULATION EXAMPLE

5.1 Simulation data and structure of the monitoring software

Behaviour of the torque signal was first studied using real data examples and references [25], [6]. On the basis of theory (section 4.1), simulation task of torque signal was divided into four linear parts. Each part had its own slope that could be altered in order to programmatically construct different process stages.

Some random variation was added to slopes so that the difference between experiment replications could be achieved. In addition, white noise was mixed to the signal. Simulation data for demonstrating the screw insertion monitoring approach was then generated based on factor levels in matrix (Table 2). Typical shapes of the torque signals are shown in Figure 6. In this example plot, altered factor is the hardness of the material.

![Figure 6. Example of the generated simulation data, three different material types.](image)

Three main classes of the normal process conditions were used as a basis for the simulations. The classes were classified according to the material type. Besides the material type, three other factors were taken into account when constructing the signal shapes.

All in all 16 different torque signals were constructed and then replicated four times with random variations. Length of the individual signal set was 100 sample points. Sample number also modelled the screw advancement during insertion process. Half of the simulation data was used for normal model parameter estimation (training) and the other half for checking and validation of the presented system. Fault data was generated by abnormally deviating the slopes in the linear parts of the signal. 24 fault data sets were
created, including five different fault types: slippage, jamming/cross-threading, no screw advancement, wrong height and missing screw.

**Figure 7.** Example of the simulated fault situation: jamming.

Monitoring system presented in Figure 3 was programmed to software at Matlab-Simulink environment so that it could be tested with simulation data sets. General flowchart of the software is presented in Figure 8.

**Figure 8.** General flowchart of the programmed software.
Each of the boxes inside the dotted box in Figure 8 is implemented as a separate function. Simulation data is read to program from an external file, 100 data points at the time. Data is then instantly processed in other modules. Initial parameters of the model were manually copied to the software from FuzzEqu /22/ development environment, since those were used only once in the beginning of the simulations.

Detailed software structure of the fault detection part is shown in Figure 9. The system includes normal condition model and five fault models. Outputs of the models are further handled and the fault is isolated at the next stage of analysis (section 5.3).

**Figure 9.** Monitoring system in Simulink environment. Input signals are analysed in separate modules. Classifier is intended for isolation of specific fault cases.

Residuals of the normal condition model are analysed simultaneously with fault models. The analysis module of the normal model output can be seen at the top of the fault models in Figure 9. Model inputs are features of measured signals, for example maximum and minimum signal values at certain points of the process. Since the input signals are derived from the same process stages on every working cycle, it is possible to construct permanent fault knowledge base in the form of fuzzy rules. In this way a heuristic knowledge can be implemented as a part of the software. Normal condition model structure and analysis method for model residuals are described in the following section.
5.2 Quality monitoring: Normal condition model

5.2.1 Model structure

Simulation data reserved for training was first analysed to determine the model structure. The autocorrelation function indicated that an autoregressive model could be suitable for this case. In a higher-order autoregressive process, the autocorrelation function is often a mixture of exponentially decreasing and damped sinusoidal components /11/. One can see the appearance of those components in Figure 10.

![Autocorrelation function of the simulation data.](image)

**Figure 10.** Autocorrelation function of the simulation data.

A partial autocorrelation function /4/ was then plotted for determination of the model order. The plot shows clear statistical significance for lags 1 and 2 (Figure 11). After the second lag, the value decreases under the 95% confidence limits (dotted line) indicating that there no statistical significant correlation at lags three and onwards.
Additional preliminary tests with different amount of lags for model inputs showed that the highest modelling performance was achieved using two lagged values of the signal. These values were chosen for final inputs. Quality monitoring of the process was carried out with linguistic equations /20/, that have been investigated for process diagnosis earlier in /19/, /21/.

Data for the normal condition model was pre-processed, consisting of four stages. First it was normalised in such a way that every individual signal data set was scaled between zero and one. After the normalisation, data points were replaced with their moving averages. Then the nonlinear scaling of the input variables was carried out. Finally, the data entering the normal condition model was rearranged to a suitable form. The nonlinear membership definitions for variables were defined using FuzzEqu development environment /22/. Figure 12 shows the final shapes of the polynomials used for scaling of inputs.

**Figure 11.** Partial autocorrelation function of the simulation data.

**Figure 12.** Membership definitions for normal condition model.
Polynomials can be seen as non-linear scaling factors describing behaviour of the data, from which they are defined. At modelling, the real values of variables are first converted using polynomials to linguistic variables (nykyinen, ykkonen and kolmonen) between −2 and 2. This action is called linguistification of variables. The resulting linguistic equation model can now be presented in the form

\[ a_1 \text{nykyinen} + a_2 \text{ykkonen} + a_3 \text{kolmonen} + b = 0, \]

where \( a_1, a_2 \) and \( a_3 \) are the parameters of the model and variable \( b \) is a bias term. After scaling the values to the same magnitude, parameter estimation is performed using a linear regression.

5.2.2 Training and validation

The model parameters in (1), defined by FuzzEqu software, were used as initial values for the further tuning of the model. Training method based on steepest descent, known also as delta rule, was applied to linguistic equation model. Parameters were adjusted recursively according to specified objective function \( E \):

\[ E = \frac{1}{2} \left( y(t) - \hat{y}(t) \right)^2, \]

where \( y(t) \) is the real value of the signal and \( \hat{y}(t) \) is the signal value predicted by the normal condition model. Value of the learning factor was set small because the initial model parameters were reasoned to be near their correct values.

After adaptation of the model (1) parameters, variable \( \text{nykyinen} \) was solved from the equation. Output of the model was then a linguistificated number, which was compared to the linguistificated value of the simulated variable \( \text{nykyinen} \). With this procedure the last stage of the linguistic modelling, delinguistification i.e. converting the linguistic values back to real values, was not necessary to implement.

Two sets of data were used in parameter estimation: training data and checking data. Training data was used for tuning of initial model parameters. Performance of the model was evaluated with checking data after every epoch. Parameters of the best training epoch, resulting a smallest error with checking data, were stored automatically and set to a final model.

The model was tested with validation data. Figure 13 shows the epochs and the errors with checking and validation data. The triangle shows the best epoch in both cases. Validation data was not used for training. It was noticed that the model performance notably improved compared to an initial one with the training method used for parameter adaptation. This indicates that linear relations between variables were not perfectly attained with non-linear scaling. Modifications to membership definitions could improve the initial model performance. On the other hand, the presented multivariate parameter estimation ensures the model performance without any tuning of polynomials. This simplifies the model identification process, especially if the updating need is small.
Figure 13. Normal model errors vs. training epochs with checking data (*) and validation data (+). Minimum error is reached at second epoch. Minimum error with validation data appears on third epoch.

The error between the normal condition model output and the actual output was calculated. In the Figure 14 one can see examples of the modelling error during simulation of normal and failed screw insertion.

Figure 14. Modelling error vs. screw advancement. (a) Successful and (b) failed screw insertion.

Figure 14 indicates that in the presence of fault, modelling error is much larger than during successful insertion. Comparison of Figure 14(a) with other simulated modelling errors didn’t show any significant differences in the magnitude of residuals.
5.2.3 Monitoring and prediction of process quality using fuzzy thresholds

It was observed that when the fault was present, modelling errors of the normal condition model had significant deviations in average or standard deviation (Figure 14). These deviations were used as symptoms to describe the current process condition. Since modelling error at non-fault situation had a normal distribution (see Appendix 1) despite of changing operation conditions in simulated process, permanent fuzzy thresholds for mean and standard deviation could be defined to handle uncertainties in these symptoms.

Unlike in /3/ where fuzzified cumulative sum and standard deviation was used as fuzzy statistical tool, the combination of mean and standard deviation was found more sensitive for monitoring the process. Fuzzy reasoning was introduced for classification between normal and faulty conditions of the system. Mean and standard deviation of the modelling error were fuzzified, and using fuzzy inference, combined to form a single value estimate of the current quality index of the screw insertion sub-system. Figure 15 presents the membership functions for the standard deviation and mean.

![Figure 15. Membership functions for mean and standard deviation of the modelling error, normal condition model.](image)

Locations and shapes of the membership functions were defined heuristically on the basis of modelling results with training data. The rule used for fuzzy inference to form a quality index is defined as follows:

(1) IF mean is <error> OR standard deviation is <error> THEN output is <error>
Output of the above rule is a single crisp index value between zero and one. It is the possibility degree of the modelling error at which it is located within defined thresholds. Interpreting the index is straightforward: values near one (OK-conclusions) mean that the process behaves normally with very high possibility, whereas values near zero indicate a strong deviation from normal behaviour. It should be noted that because of small margins in limits, even the index value is zero the process may still function at least some degree. Fuzzification of variables makes possible to use uncertain statistical information. It also enables a continuous index value.

The quality index of normal condition can be now employed for quality monitoring of the screw insertion process. Example is shown in Figure 16, where the trend of the index value is presented as a function of successful screw insertions.

![Figure 16](image)

**Figure 16.** Example of utilising the quality index for monitoring of screw insertions.

If the index can be regarded normally distributed in a non-fault situation, the additional use of index value along with basic control charts and tests for trends in time series are justified. Advantage of this implementation is that both mean and standard deviation can be joined to a single value instead of monitoring each of them separately. It can help to detect deviations from normal conditions more reliable and faster. Observation of multiple statistical quantities at the same time may complicate the interpretation of the results.

Normal condition model of the process is an essential part of the monitoring system. Fault models are another important component. These models are needed for all faults supposed to be detected and isolated.
5.3 Fault detection and isolation: Fault models

Faults to be detected were chosen with the help of failure modes and effects analysis. The most common faults of screw insertion sub-system included for example jamming (Figure 7). Expert knowledge was utilised to fault detection task, it was represented in software with fuzzy rules. Fuzzified values of measurement signals were used in fuzzy reasoning. A fuzzy rule base was developed for five different fault types.

The possibility of certain fault type was estimated by using fuzzy inference; fuzzy rule-firing example is presented in Figure 17. The output is a possibility value 0.822 for this type of fault example. The related fuzzy rule is presented below:

(1) IF p1 is <high> AND p2 is <low> AND p3 <low> THEN output is <error> ,

where p1, p2 and p3 are fuzzified parameters that approximate the their crisp values at different levels of measured signals.

Few of the fault models were designed to detect only one fault, whereas other models were used to detect a group of faults. In the first case the fault was isolated directly. In the latter case the fault model output had to be analysed further to isolate the fault. Certain types of faults were more reliably detected when the signal time series was divided in several parts, and each part was then modelled locally.

The possibilities given by the fault models and the normal condition model were compared to determine the current process condition. The greatest possibility of the models output was chosen. The normal condition model was used to ensure the final conclusion because it was found reliable in performed tests. The following general rules were applied to fault isolation module of the system:

(1) IF fault model output is <ok> AND normal condition model output is <ok> THEN process condition is <ok>
(2) IF fault model output is <ok> AND normal condition model is <fault> THEN process condition is <unknown fault>
(3) IF fault model output is <fault> THEN process condition is <fault>

The developed rule base was tested with simulation data containing both ok- and faulty signals. The estimate of the process condition was given verbally. The list also included the truth-value of the model conclusion, which is the possibility estimate of the process condition (Figure 18).

![Figure 18. Example of a verbal process condition list during multiple screw insertions.]

In some cases two possibility indices are presented at the same time. For example in case 15 this implies that a fault is present with possibility of 0.95 and the type is jamming with possibility of 0.67. If possibilities for all modelled situations were small, process condition was reasoned to be an unknown fault (case 5, Figure 18). Further analysis of the case 5 showed that the random variations had affected to a signal making it dissimilar with any predefined faults.

Detection of unknown process conditions was included in the system for future development of case based reasoning (CBR) part. This method could learn unknown abnormal events by recording and training itself with the help of human operator and data stored to database.
5.4 Comparative study: principal components and neural networks

Monitoring of the screw insertion sub-system was also tested with combination of principal components (PC) and neural net. The main idea was to use PC-analysis to extract highly uncorrelated inputs for neural net. Another motivation for the use of PCs was their known capability to be able to reduce the dimension of the input data. This could result fewer inputs leading to a simple structure of the system. Details of the PC-analysis are discussed in /17/. PCs have been used as input variables for example in /26/ to construct neural net models of the molecular beam epitaxy process.

Principal components were formed making good use of Singular Value Decomposition function in Matlab-environment. Training, testing and fault data values over screw advancement were transformed to a new set of values, namely principal components. Data sets of successful insertion (training data) were used to estimate the parameters of the eigenvectors. Defined parameters were then applied to transform other data sets into PCs. Two first PCs explained 78% of the variance in training data. Components were plotted against each other to check if there are any similarities between individual screw insertions (Figure 19). It can be seen that there is a significant difference in positions of faulty and successful insertions within the plot coordinate.

![Figure 19. Two first principal components of the simulation data, PC2 vs. PC1. Solid line divides the area to successful and failed insertions.](image-url)
Various neural net structures and different amount of PCs as input variables were tested for monitoring task. Best results were obtained using four first components that explained over 90% of the variance in training data. Net structure was a 1-hidden-layer feed-forward backprop, trained with Levenberg-Marquardt algorithm. Activation function was a hyperbolic tangent sigmoid. Optimised model structure contained eight neurons in a hidden layer and bias terms in hidden and output layer. Neural net was trained to recognise successful insertions so that output of the net was supposed to be zero. A separate vector of zero values was defined to target values during learning. Several runs with different starting points and initial network weights were performed. After the training, model was tested with testing data set and fault data. Results are presented in Figure 1.

![Graph of neural network output](image)

Modelling results of the neural net trained with first four principal components as inputs. Values near zero indicate successful insertion. Large deviations from zero with fault data (solid line) indicate correct classification.

Figure 1 shows that the net accurately recognises successful insertions of the training data set, i.e. the dotted line equals zero. Model is also quite insensitive to differences in testing data of normal conditions (dashed line). However, the model doesn’t notice several fault situations when tested with fault data. Value of the model output should deviate from zero if the fault occurs. This means that the solid line describing the output with fault data should have values far away from zero in every case. Using modelling error with testing data as threshold, eight out of 24 cases were misclassified with presented method. From the Figure 1 it can be seen also that neural net model output exposes the deterministic
pattern of simulation data, when introduced with fault situations. This kind of information may be useful in further analysis of process conditions.

A partial explanation for classification error of faults can be seen in Figure 19. In the plot coordinate, distance between normal and fault insertion is very small in some cases. On that area (near the solid line), the trained neural network is sensitive to even small deviations in the values of principal components. This complicates the mildly non-linear classification of certain insertions with tested model structure.

The small amount of training data doesn’t favour presented approach. Early stopping training strategy could be applied and tested with more complicated model structure. Further studies are also needed to investigate other network structures more thorough. It is concluded that monitoring with principal components and neural net has the potential for monitoring task in this case.
6 CONCLUSIONS

In this work a general monitoring framework is proposed for quality monitoring and fault detection of automated manufacturing cells. Details of the monitoring system structure are discussed and programmed to software. Principles of the monitoring approach are demonstrated with a simulation example.

In the presented framework, symptoms to be observed are defined with the help of experts and failure analysis. Models used for monitoring are based on systematically collected data and expert knowledge. Experimental design produces both data for optimisation of machine parameters and representative training data for modelling. This procedure enables robust models and adaptation to new machines and processes. Trained with systematically collected experimental data, the model is guaranteed to include behaviour of the system over wide range of operation conditions. On the other hand, expert knowledge can be fully utilised by using rule-based models where suitable.

The use of model bank structure at the monitoring system allows specific models for each situation. Since the models are separate, the structure of the system is flexible whereas outputs of the model bank are analysed jointly resulting a permanent module. This kind of an open system can be modified to meet the requirements of other assembly cells.

Rather than run-by-run quality monitoring, the use of autoregressive linguistic time series model makes possible for example the continuous observation of the screw insertion process. It also enables the collection of more specific information about situations between the beginning and end of the individual working cycle. Training with data obtained via designed experiments, time series model needs the minimum amount of the data compared to run-by-run methods. Simulations show that the normal condition model is capable to generalise when introduced with unseen validation data. Results indicate that the developed model is robust against uncertainties and noise in the data. The normal condition model inside the model bank can additionally fine-tune itself using delta rule algorithm for parameter adaptation.

Fuzzy thresholds were developed for instant change detection in normal condition model residuals. The fuzzified values of error average and standard deviations were put in the form of simple fuzzy rule, resulting a statistical quality control tool. Combination of these variables is a single quality index of the process. It was found reliable despite of uncertainties in applied statistical variables. Prediction of quality and degradation of tool wear can be considered using the constructed index together with basic statistical tools.

Fault models were presented in the form of fuzzy rule base by utilising expert knowledge. Rule base consists of causal relations between signals to be measured. Fault isolation was performed analysing the outputs of the rule base. Conclusions of the fault models during simulations were reported automatically in the form of failure list that also included the possibility index for every occurred fault.
It should be noted that the simulation data generated in this work only suits to be used for demonstration of the presented methods. Data presents a rough approximation about measurement signals in screw insertion process. In addition, prediction of tool wear using presented modelling approach is not presented due to lack of real data.

Unknown response of the monitoring system can be further employed for learning purposes. If the monitoring system produces an unknown fault output, case based reasoning can be introduced so that next time when the fault occurs, it is automatically identified using the former experience. Database of the observed symptoms in fault situations is required.

Presented framework is intended for quality monitoring and early detection of failures at electronics manufacturing systems, especially at ramp-up stage. Fault detection and isolation may be further utilised for diagnosis i.e. identification of the fault causes. After detection of process conditions, the main objective is to recover from failure and resume productivity together with quality. Future work will focus on prediction of the process quality parameters and for preventive fault diagnosis, which is a necessary link between process monitoring and fault recovery operations.
REFERENCES


/39/ Seneviratne, L., D. and Visuwan, P.: Weightless Neural Network Based Monitoring of Screw Fastenings in Automated Assembly, Neural Information


Weber screwdriver systems Inc., 45 Kensico Drive, Mount Kisco, USA.

Test for normality of errors of the AR(2) -linguistic equation model: Average of modelling error per screw insertion, 48 cases. (Calculated with DATAPLOT-software)

ANDERSON DARLING 1-SAMPLE TEST
THAT THE DATA COME FROM A NORMAL DISTRIBUTION

1. STATISTICS:
   NUMBER OF OBSERVATIONS = 48
   LOCATION PARAMETER = 0.2300554E-02
   SCALE PARAMETER = 0.1136515E-02

   ANDERSON DARLING TEST STATISTIC VALUE = .3948022

2. CRITICAL VALUES:
   90 % POINT = 1.062000
   95 % POINT = 1.321000
   97.5 % POINT = 1.591000
   99 % POINT = 1.959000

3. CONCLUSION (AT THE 5% LEVEL):
   THE DATA DO COME FROM A NORMAL DISTRIBUTION