IIE Transactions (2009) **41**, 615–625
Copyright © "IIE"
ISSN: 0740-817X print / 1545-8830 online
DOI: 10.1080/07408170802389316

Multiscale mapping of aggregated signal features to embedded time–frequency localized operations using wavelets

JIONGHUA JIN[1,*] and JING LI[2]

[1] *Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor MI 48109-2117, USA*
E-mail: jhjin@umich.edu
[2] *Department of Industrial Engineering, Arizona State University, Tempe, AZ 85287-5906, USA*

Received May 2007 and accepted May 2008

Aggregated signals are referred to as the measurements of system-level responses generated by the multiple operations embedded in a system. While an extensive literature exists on the analysis of general signal profiles, limited research has been performed on the topic of how to map features of the aggregated signals to the responses of individual operations, which is important for individual operation performance monitoring and assessment. In this paper, a two-step mapping algorithm is developed to obtain those mapping features using a mutiscale wavelet analysis integrated with statistical hypothesis testing and engineering knowledge. It is shown that multiscale wavelet analysis is effective for mapping aggregated signals to the embedded individual operations that generate localized time–frequency responses. This algorithm is further demonstrated in a stamping process, in which the extracted wavelet coefficients of aggregated press tonnage signals are explicitly mapped to individual or a few contributing embedded operations. The mapping allows for efficient monitoring and quality assessment of the embedded operations based on the aggregated signals, thereby avoiding installing additional in-die sensors in all operations.

[Supplementary materials are available for this article. Go to the publisher's online edition of *IIE Transactions* for the following free supplemental resource: Appendix]

Keywords: Aggregated signal, profile analysis, wavelet analysis, multistage process

1. Introduction

In general, a complex system consists of multiple embedded operations. Although the separate measurement of individual operation responses can provide explicit diagnostic information for system monitoring, it is often not possible or economic to install individual sensors at each operation. In practice, a commonly used method is to measure the combined response of embedded operations; this combined response is called an aggregated signal. Therefore, it is desirable to know whether it is possible to monitor the performance of individual operations based on available aggregated signals.

For example, in a transfer/progressive stamping or forging process, strain gage sensors are usually installed on press machines to measure the total press tonnage force (i.e., an aggregated signal), which is comprised of the forces of all embedded operations. In contrast, in-die sensors, which are installed inside the dies of individual operations to measure their respective forces, are very rarely used in real-world production, because of the extra costs for purchasing, installing, inspecting, and replacing the sensors, as well as the accompanying complexity of die changes. Similarly, in a computer network, a border router is usually available to collect aggregated packet traffic data (i.e., an aggregated signal) corresponding to the total number of queries from all relevant nodes in the network, while the number of queries from individual nodes is not separately measurable. A third example is a transportation network in an urban area, where traffic flow is usually measured at certain main intersections (i.e., an aggregated signal), but not along each branch road whose traffic flow passes through the intersection and which contributes to its overall traffic flow. In these three examples, it is highly desirable to have the ability to use aggregated signals to monitor and detect abnormalities in individual manufacturing operations, network nodes and branch road traffic flows, for the purpose of quality assurance in stamping and forging, computer network intrusion detection and better route planning and traffic control,

*Corresponding author

0740-817X © 2009 "IIE"

respectively. There is a substantial need for research on how to map features of system-level aggregated signals to the responses of embedded operations. The advantage of using such mapping features is improved diagnostic capabilities for system monitoring, which can be shown in two ways: first, when there is a one-to-one mapping between a feature and an individual operation; this feature can be used to monitor the operation exclusively. Second, even for some operations lacking dedicated features, an explicit mapping between a feature and some of the operations, rather than all of the embedded operations, can expedite the search for the specific failures in operations.

In this research, we focus on systems consisting of multiple embedded operations, each of which generates a localized "time–frequency" response. In other words, in order to allow multiple embedded operations to be monitored separately based on the aggregated signals, these operations must engage in the system at different time segments (i.e., possibly overlapping but not identical time periods) and/or have their energy concentrated within overlapping but not identical frequency bands. This characteristic is shared by the systems in the previous three examples. Specifically, past research in stamping and forging process control (Jin, 2004; Jin and Shi, 2005) has found that the actual working range of each individual operation only occupies a limited, mostly different, portion of the entire time cycle of the aggregated signal, and each embedded operation has its energy concentrated in mostly different frequency bands. Likewise, the number of queries from the individual nodes of a computer network also forms localized time–frequency signals due to the differences among the individual nodes in the network in terms of usage time and usage patterns (Barford *et al.*, 2002). Similar phenomena can be observed in the traffic flow passing through an intersection from branch roads (Guo and Jin, 2006).

An extensive literature exists on the analysis of signal profiles and the extraction of features for system monitoring and diagnoses (Woodall *et al.*, 2004). Generally, multivariate statistical methods, such as Principle Component Analysis (PCA) (Johnson and Wichern, 2002), PCA-based pattern recognition (Ceglarek and Shi, 1996), factor analysis (Apley and Shi, 2001; Johnson and Wichern, 2002), variance component analysis (Rao and Kleffe, 1988; Zhou *et al.*, 2003) and signature metrics approaches (Kibarian and Strojwas, 1991; Gardner *et al.*, 1997), focus on analyzing signals in the time domain. Fourier transforms (Bracewell, 2000) decompose signals in the frequency domain, but these are applicable only to stationary signals and not to localized time–frequency non-stationary signals. Short-time Fourier transforms can be used to analyze localized time–frequency signals but at a fixed resolution (or scale) in both the time and frequency domains. In contrast, wavelet analysis (Mallat, 1989) is able to decompose a signal at multiscale time–frequency domains, and the resulting wavelet coefficients capture localized signal features at different time segments and different frequency bands.

Most past research on the application of wavelet analysis to system monitoring and diagnosis resides in the general field of profile or functional data monitoring, which aims to develop control charts for different types of profiles, including linear (Stover and Brill, 1998; Kang and Albin, 2000; Kim *et al.*, 2003; Mahmoud and Woodall, 2004; Mahmoud *et al.*, 2007) or non-linear profiles (Young *et al.*, 1999; Walker and Wright, 2002; Ding *et al.*, 2006), profiles that can be represented in a parametric (Stover and Brill, 1998; Kang and Albin, 2000; Kim *et al.*, 2003; Mahmoud and Woodall, 2004; Mahmoud *et al.*, 2007) or non-parametric (Winistorfer *et al.*, 1996; Gardner *et al.*, 1997; Fan and Lin, 1998) forms, and complex waveform profiles that have rich information in a localized time domain (Jin, 2004), or in both localized time and frequency domains (Jin and Shi, 1999, 2001; Lada *et al.*, 2002). In this area of research, wavelet analysis is often used for signal denoising, profile fitting or feature-preserving data compression (Jin and Shi, 1999).Very little research, however, has been published on the application of wavelet analysis to the profile analysis of aggregated signals for embedded operation monitoring. This opens a new research area on how to map features (i.e., the wavelet coefficients) of the aggregated signals to the embedded individual operations.

This paper uses a multistage progressive stamping process to illustrate the development of a data analysis method for exacting features of aggregated signals that can be mapped to the corresponding embedded operations. In the literature of stamping signal analysis, the Discrete Wavelet Transform (DWT) approach has been successfully used for multiscale feature extraction (Jin and Shi, 1999). Therefore, this paper will also apply the DWT approach to the aggregated stamping force signals in the first step of data analysis. Then, the extraction of mapping features is performed by identifying the wavelet coefficients corresponding to each individual operation.

The remainder of this paper is organized as follows: Section 2 introduces data collection and the characteristics of the collected signals. Section 3 presents the DWT approach for the decomposition of the collected signals into wavelet coefficients and Section 4 proposes a method to map the wavelet coefficients of the aggregated signals to the corresponding embedded individual operations. Section 5 provides a case study and finally conclusions are drawn in Section 6.

2. Data collection and signal characteristics in multistage progressive stamping processes

Two types of signals are collected in order to establish the mapping relationship between the aggregated signals and the individual contributing operations embedded in a progressive stamping process. One is the aggregated total tonnage signals, which can be collected continuously during production. The other is the individual operation tonnage

force signals, which are collected offline using specially designed physical experiments. This section introduces the progressive stamping process, the collection of these two types of signals and the signal characteristics.

2.1. *Online collection of aggregated stamping tonnage signals*

A progressive stamping process requires several sequential operations to produce one part, hence the use of the term "multistage" process. This process adopts a progressive die in which the sequential operations can be performed at different stations through multiple embedded dies. During this process, an automatic feeding device pushes the raw material strip through all of these stations by advancing the strip to the next station at each press stroke. At each station, one or more operations are performed, and at the final station, a finished stamping part is formed. The produced part is expected to meet the design specifications on both shape and dimension. For example, Fig. 1 shows a progressive stamping process for producing doorknobs. The process is equipped with a progressive die having six stations. At these stations, the following operations are sequentially performed: notch and cutoff (together at one station); blanking; draw; redraw; second redraw; and bulging.

In a progressive stamping process, strain gage sensors are usually installed on the linkages or columns of a stamping press. These are called "press sensors," and they are used to collect the aggregated total press tonnage force signals that reflect the overall performance of all of the embedded individual operations. For example, in Fig. 1, four press sensors are installed on the stamping press (one on each press column), and the summation of the signals measured by these four sensors is an aggregated signal contributed to by all seven individual operations. One sample of the aggregated signal is given in the top-right panel of Fig. 1.

In this paper, the aggregated tonnage signal is denoted by a random vector , $\mathbf{F} = [F_1, \ldots, F_n]^{\mathrm{T}}$, where n is the number of data points measured within the cycle of a press stroke (i.e., from the time that the press moves down to the die, through the time of the executions of the operations, to the time that the press is lifted up) with each data point corresponding to a sampling time unit. This aggregated signal is the summation of the individual operation tonnage forces plus the initial support force of the cushion that is located under the lower bed of the progressive die, as shown in Fig. 1. Unlike the aggregated signal, the signals for these individual operation tonnage forces and the initial cushion force are not measurable during continuous production. They are denoted by $\tilde{\mathbf{F}}^i = [\tilde{F}_1^i, \ldots, \tilde{F}_n^i]^{\mathrm{T}} (i = 0, \ldots, m)$, where m is the number of operations, $i = 0$ represents cushion, and "$\sim$" implies that the signals are not measurable during continuous production. Therefore, the physical relationship between $\mathbf{F}$ and $\tilde{\mathbf{F}}^i$ is

$$\mathbf{F} = \sum_{i=0}^{m} \tilde{\mathbf{F}}^i. \tag{1}$$

2.2. *Offline collection of individual operation signals through physical experiments*

Although the individual operation signals are not measurable during continuous production, they can be measured using the press sensors through a designed offline experiment that has been proposed by Jin and Shi (2005). In this paper, the individual operation signals obtained from the offline experiment are denoted by random vectors $\mathbf{F}^i = [F_1^i, \ldots, F_n^i]^{\mathrm{T}}$ $(i = 0, \ldots, m)$. Figure 2 shows one example of the individual operation signals $\mathbf{F}^i$ $(i = 0, \ldots, 7)$,

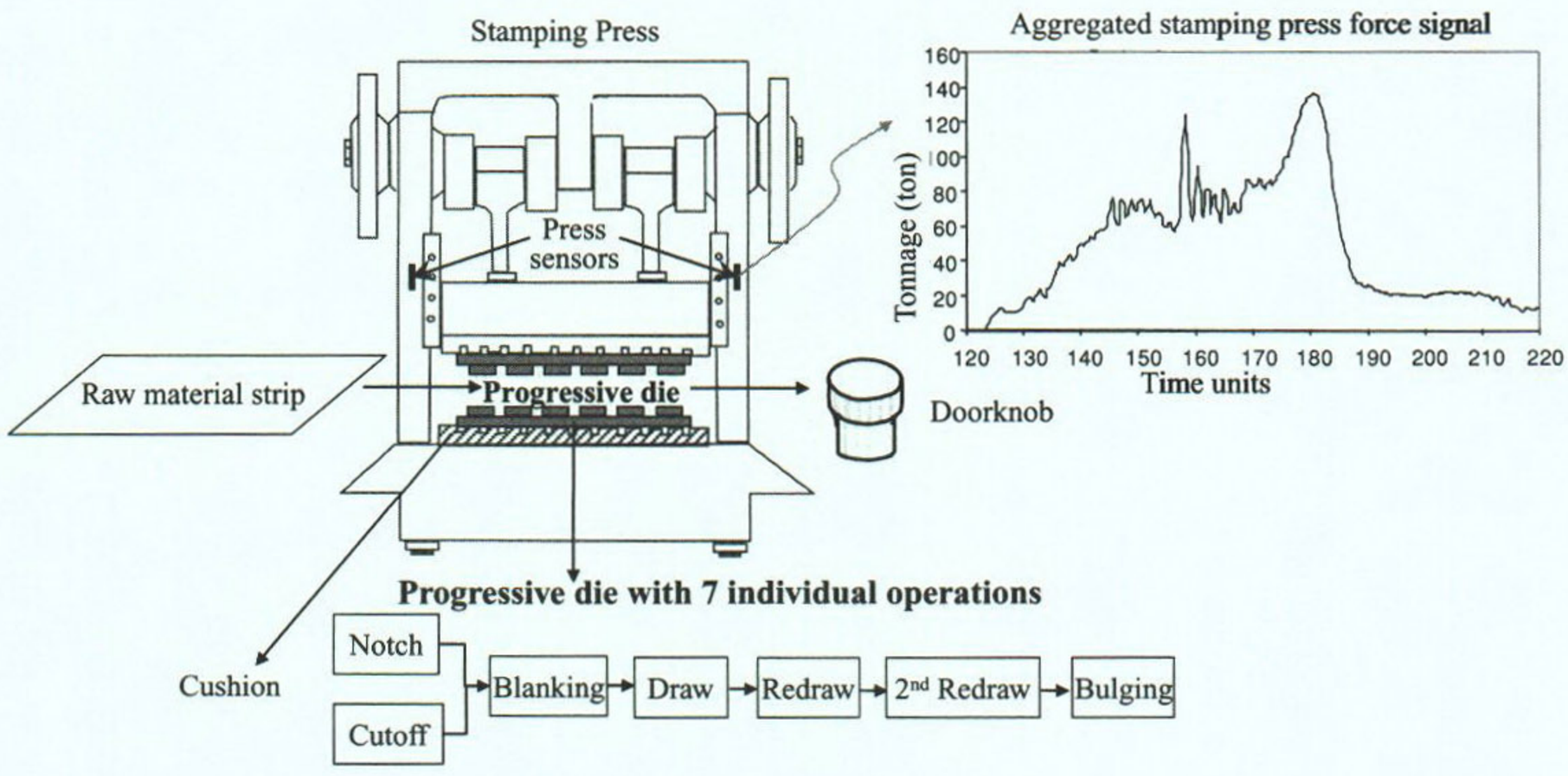

Fig. 1. A multistage progressive stamping process for doorknob production.

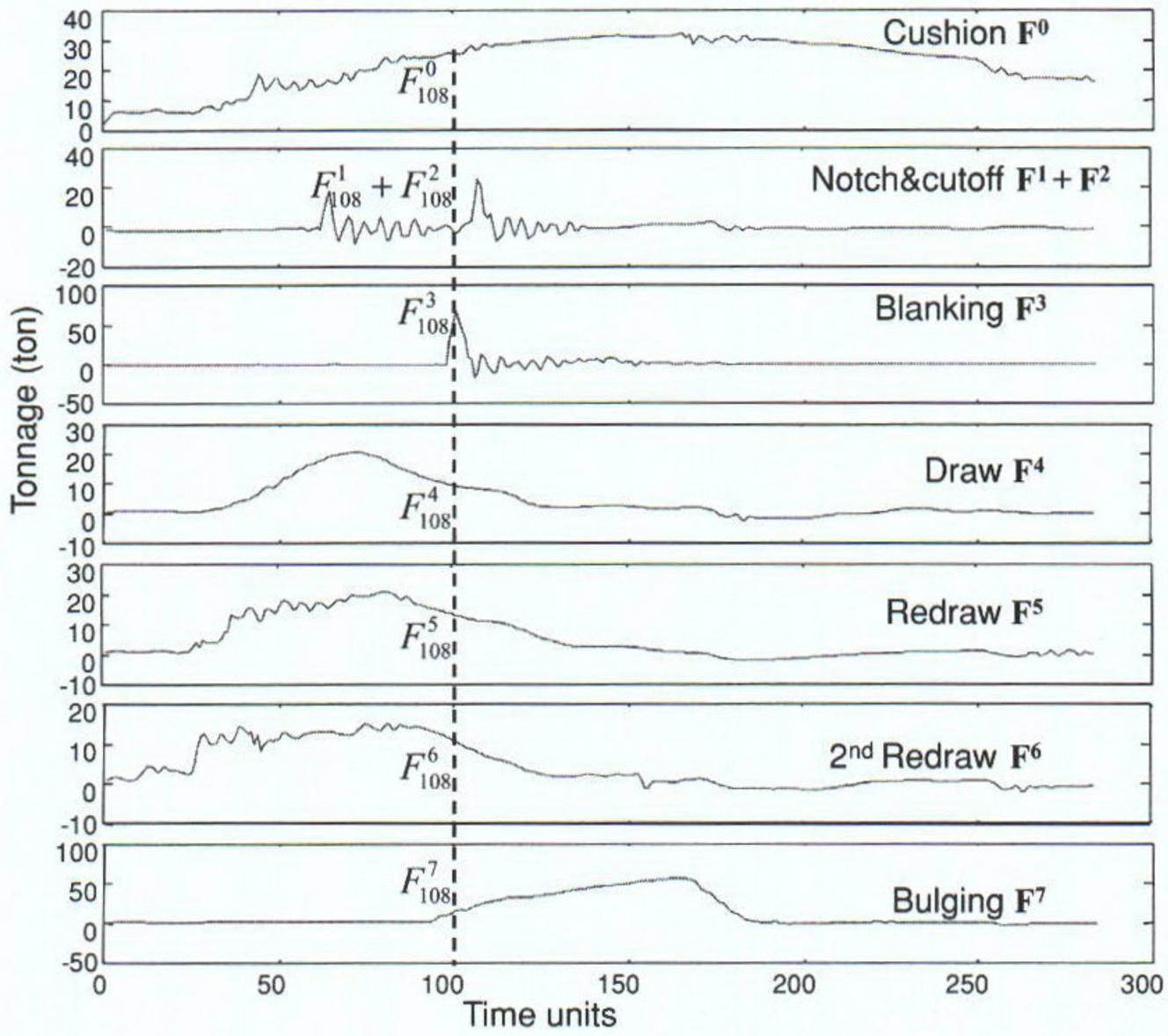

Fig. 2. An example of the individual operation signals $\mathbf{F}^i$ $(i = 0, \ldots, 7)$ obtained from the offline experiment at the stamping process in Fig. 1.

collected in the experiment. Furthermore, it is reasonable to assume that the mean vector of $\mathbf{F}^i$, $\boldsymbol{\mu}_{\mathbf{F}^i}$, is equal to that of $\tilde{\mathbf{F}}^i$, $\boldsymbol{\mu}_{\tilde{\mathbf{F}}^i}$ under the same working conditions, that is

$$\boldsymbol{\mu}_{\mathbf{F}^i} = \boldsymbol{\mu}_{\tilde{\mathbf{F}}^i}. \tag{2}$$

By combining Equations (2) and (1), the mean vector of the aggregated signal $\mathbf{F}$, $\boldsymbol{\mu}_{\mathbf{F}}$, can be expressed as

$$\boldsymbol{\mu}_{\mathbf{F}} = \sum_{i=0}^{m} \boldsymbol{\mu}_{\mathbf{F}^i}. \tag{3}$$

In order to map features of the aggregated signal to the embedded individual operations, aggregated signals are also collected from continuous production under the same working conditions as the offline experiment. Figure 3 shows one example of the aggregated signals. When comparing the aggregated signal $\mathbf{F}$ in Fig. 3 and the individual operation signals $\mathbf{F}^i$ in Fig. 2, it is important to note that searching for the mapping features only in the time domain of the aggregated signal is possible only to a very limited extent for each individual operation because the aggregated tonnage force at a time unit k may be influenced by the corresponding tonnage forces at many or all of the individual operations. For example, the peak tonnage force in the aggregated signal in Fig. 3, F_{108}, is influenced by the tonnage forces at all seven operations, F_{108}^i, $i = 1, \ldots, 7$, and the tonnage force of the cushion, F_{108}^0; that is, $\mu_{F_{108}} = \sum_{i=0}^{7} \mu_{F_{108}^i}$ according to Equation (3). Thus, monitoring F_{108} is ineffective for assessing individual operation performance.

While a time-domain analysis of the aggregated signal alone may not be adequate to identify the mapping features for individual operations, an integrated analysis that fully considers the rich local information of the signal in both the time and frequency domains can serve this purpose effectively. This time–frequency analysis can be achieved by the use of a DWT, which will be introduced in the next section.

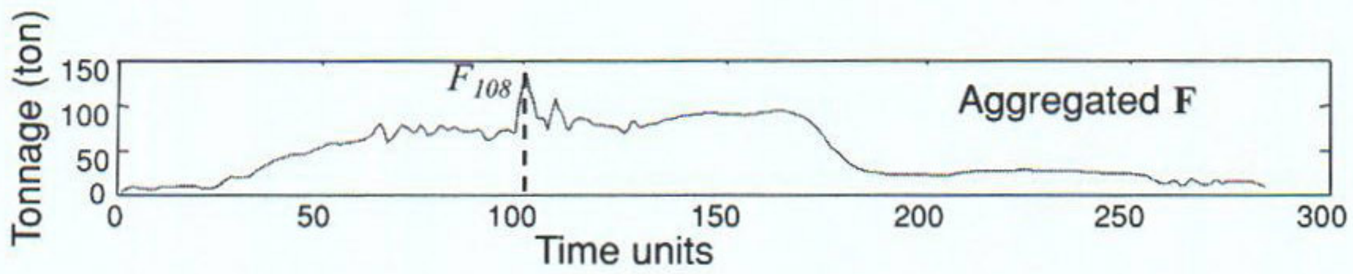

Fig. 3. An example of the aggregated signals $\mathbf{F}$ obtained online from continuous production at the stamping process in Fig. 1.

3. Signal decomposition based on a DWT

Past research on stamping processes provides abundant evidence that individual operations are reflected in a *localized* time–frequency domain of the aggregated tonnage signal. In the time domain, Jin (2004) and Jin and Shi (2005) found that the working range of each operation is only a limited portion, called a segment, of the entire cycle of the aggregated signal. In the frequency domain, it is known from physical design knowledge that the energy of each operation is concentrated in a certain frequency range. For example, notch, cutoff and blanking operations correspond to a high-frequency range, but draw and bulging operations correspond to a low-frequency range. Therefore, the significant advantage of using a DWT is that the multiscale wavelet coefficients of the aggregated signal are localized not only in the time domain, but also in the frequency domains and this leads to an additional separability among the operations at different scales. A brief review of the DWT approach can be found in Li *et al.* (2007).

In this paper, a DWT is performed on the aggregated signal $\mathbf{F}$ by taking the product of a DWT matrix, $\mathbf{W}$, and $\mathbf{F}$, thus producing a set of wavelet coefficients, $\boldsymbol{\theta}$, that is

$$\boldsymbol{\theta} = \mathbf{WF}. \tag{4}$$

Here, $\mathbf{W}$ is determined by choosing a particular wavelet basis and a maximum decomposition level d. Because a wavelet basis has a finite length, the wavelet coefficients $\boldsymbol{\theta}$ are localized in the time domain. Moreover, the composition of the wavelet coefficients vector $\boldsymbol{\theta}$ is $\boldsymbol{\theta} = [\boldsymbol{\theta}'^{\mathrm{T}}_d, \boldsymbol{\theta}^{\mathrm{T}}_d, \ldots, \boldsymbol{\theta}^{\mathrm{T}}_j, \ldots, \boldsymbol{\theta}^{\mathrm{T}}_1]^{\mathrm{T}}$, where $\boldsymbol{\theta}_j = (\theta_{j,1}, \ldots, \theta_{j,n_j}, \ldots, \theta_{j,N_j})^{\mathrm{T}}$, $j = 1, \ldots, d, d'$, $\boldsymbol{\theta}_j$ $(j = 1, \ldots, d)$ are called the detail coefficients at decomposition level j, and $\boldsymbol{\theta}_{d'}$ are called the approximation coefficients at decomposition level d. Because each decomposition level corresponds to a certain frequency range, the wavelet coefficients at each level are localized in the frequency domain.

To map the wavelet coefficients of the aggregated signal to individual operations, the same DWT is also applied to the individual operation signals collected from the offline experiment in Section 3, $\mathbf{F}^i$, and the corresponding wavelet coefficients, $\boldsymbol{\theta}^i$, $i = 0, \ldots, m$, are obtained, that is

$$\boldsymbol{\theta}^i = \mathbf{WF}^i, \tag{5}$$

where $\boldsymbol{\theta}^i = [(\boldsymbol{\theta}^i_{d'})^{\mathrm{T}}, (\boldsymbol{\theta}^i_d)^{\mathrm{T}}, \ldots, (\boldsymbol{\theta}^i_j)^{\mathrm{T}}, \ldots, (\boldsymbol{\theta}^i_1)^{\mathrm{T}}]^{\mathrm{T}}$ and $\boldsymbol{\theta}^i_j = (\theta^i_{j,1}, \ldots, \theta^i_{j,n_j}, \ldots, \theta^i_{j,N_j})^{\mathrm{T}}$.

Furthermore, because of Equation (3), the relationship between a wavelet coefficient of the aggregated signal, θ_{j,n_j}, and the corresponding wavelet coefficients of individual operation signals, θ^i_{j,n_j}, can be expressed as

$$\mu_{\theta_{j,n_j}} = \sum_{i=0}^{m} \mu_{\theta^i_{j,n_j}}.$$

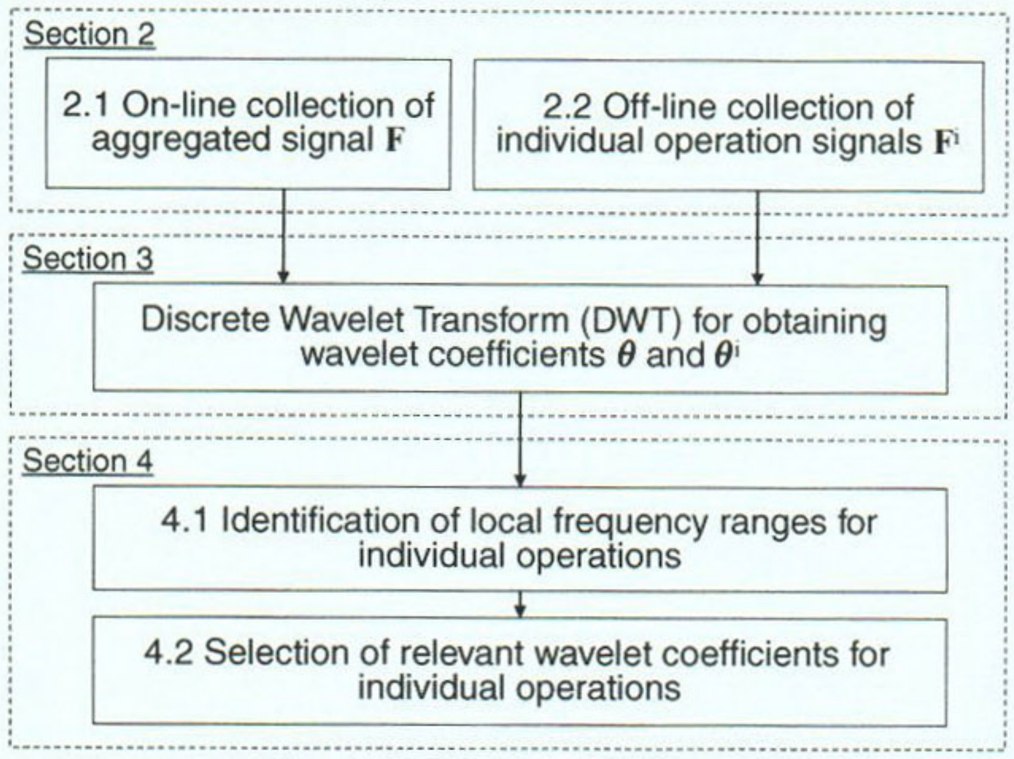

Fig. 4. Three fundamental components of the method for mapping aggregated signal features to embedded operations.

4. Mapping wavelet coefficients of an aggregated signal to embedded operations

As shown in Fig. 4, there are three fundamental components in the method for mapping aggregated signal features to the embedded contributing operations. These include the collection of signals, discussed in Section 2; the DWT of the signals, discussed in Section 3; and a two-step mapping algorithm, which will be discussed here in Section 4.

The first step of this mapping algorithm is to identify the local frequency range of each individual operation. This is achieved by applying a power spectrum analysis to the wavelet coefficients of the individual operation signals collected by offline experimental tests, i.e., $\boldsymbol{\theta}^i$, and identifying the energy-concentrated wavelet decomposition levels. Then, for each energy-concentrated level of an operation, the second step of the mapping algorithm is to select those wavelet coefficients at this level that are relevant to this operation in the time domain. The second-step selection is physically justified by the fact that an individual station in the progressive die is engaged only at a particular local time segment of a press stroke cycle; thus wavelet coefficients falling outside the time segment are not relevant to the operation performed at this station.

4.1. *Identification of local frequency ranges for individual operations*

It is known from physical design knowledge that each individual operation may generate signals localized within a certain frequency range. For example, notch, cutoff and blanking are high-frequency operations, but draw and bulging are low-frequency operations. After the frequency range of an operation is identified, only those wavelet coefficients of the aggregated signal that fall into this frequency range can be used to assess the performance of this corresponding operation.

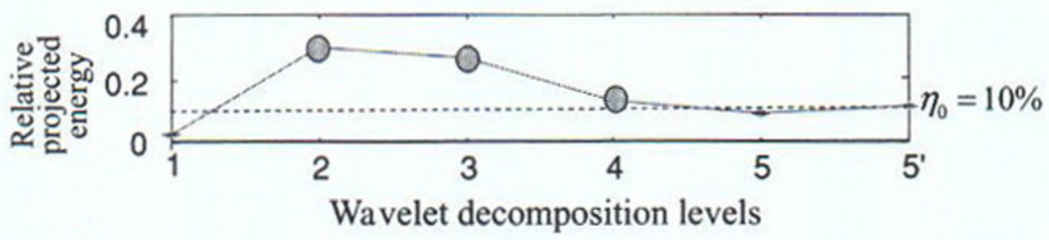

Fig. 5. Relative projected energy at levels 1 through 5 for blanking operation, with levels 2, 3 and 4 being the energy-concentrated levels for this operation.

There is a correspondence between frequency ranges and wavelet decomposition levels. Thus, the identification of the possible frequency range for each individual operation can be translated into a search for the wavelet decomposition levels on which the energy of an operation concentrates. These levels, called energy-concentrated levels, can be identified through a wavelet power spectrum analysis. In this analysis, first, the total energy of an individual operation signal $\mathbf{F}^i$ is projected onto different wavelet decomposition levels. Specifically, the projected energy at level $j \in \{1, \ldots, d, d'\}$, denoted by H^i_j, is computed by an inner product of all the wavelet coefficients at this level, i.e., $\mathrm{H}^i_j = (\boldsymbol{\theta}^i_j)^{\mathrm{T}} \boldsymbol{\theta}^i_j$. The relative projected energy at level j is defined as the projected energy H^i_j normalized by the total energy of the signal, that is

$$ {}^{\mathrm{R}}\mathrm{H}^i_j = \frac{(\boldsymbol{\theta}^i_j)^{\mathrm{T}} \boldsymbol{\theta}^i_j}{\sum_{j=1}^{d'} (\boldsymbol{\theta}^i_j)^{\mathrm{T}} \boldsymbol{\theta}^i_j}. $$

Next, the energy-concentrated level, denoted by $\mathbf{r}^i$, is all the levels with a relative projected energy larger than a certain threshold η_0, that is

$$ \mathbf{r}^i = \left\{ j | {}^{\mathrm{R}}\mathrm{H}^i_j > \eta_0, j \in (1, \ldots, d, d') \right\}, \tag{6} $$

where η_0 is the noise energy level due to natural process variation and $\eta_0 = 10\%$ for a typical stamping process. For example, Fig. 5 gives a plot of the relative projected energy at levels $j \in \{1, \ldots, d, d'\}$ for the "blanking" operation ($i = 3$) of the stamping process in Fig. 1. The energy-concentrated levels for blanking are identified as levels 2, 3, and 4, i.e., $\mathbf{r}^3 = \{2, 3, 4\}$.

Through the power spectrum analysis, we can divide the wavelet coefficients of the aggregated signal into two categories for each individual operation i: those that belong to the energy-concentrated levels of operation i, i.e., $\boldsymbol{\theta}_{\mathbf{r}^i} = \{\boldsymbol{\theta}_j | j \in \mathbf{r}^i\}$, and those that do not, i.e., $\boldsymbol{\theta} \backslash \boldsymbol{\theta}_{\mathbf{r}^i}$ (the complement of $\boldsymbol{\theta}_{\mathbf{r}^i}$ with respect to $\boldsymbol{\theta}$). It is physically justified that the wavelet coefficients in $\boldsymbol{\theta} \backslash \boldsymbol{\theta}_{\mathbf{r}^i}$ must be irrelevant to operation i and $\boldsymbol{\theta}_{\mathbf{r}^i}$ contains all possibly relevant wavelet coefficients. Furthermore, considering the time-domain localization property of each individual operation, not all coefficients in $\boldsymbol{\theta}_{\mathbf{r}^i}$ are relevant to operation i. Therefore, there is a need to further formulate a procedure to select the truly relevant coefficients from $\boldsymbol{\theta}_{\mathbf{r}^i}$, which will be discussed in the next section.

4.2. *Selection of relevant wavelet coefficients for individual operations in the time domain*

Past stamping research has revealed that each individual stamping operation has a localized working range in the time domain. Therefore, even if a wavelet coefficient of the aggregated signal is at an energy-concentrated level of an operation, the coefficient may not necessarily be relevant to this operation unless the time segment covered by this coefficient is within this operation's working range. For example, Fig. 6 plots the wavelet coefficients of the aggregated signal at level 2, which is an energy-concentrated level of blanking. To determine whether all these coefficients are relevant to blanking, the wavelet coefficients of the offline collected blanking signal are also plotted. It can be seen that the two sets of wavelet coefficients closely match within the interval indicated by the dash-lined box, which corresponds to the working range of the blanking operation, which is known from engineering design.

The problem of selecting relevant wavelet coefficients of the aggregated signal of an embedded operation can be equivalently solved by mapping each wavelet coefficient to its "contributing" operations. For a wavelet coefficient θ_{j,n_j}, the contributing operations are defined as a set of operations $\mathbf{M}(\theta_{j,n_j}) \subseteq \{0, \ldots, m\}$ such that:

$$ \mu_{\theta_{j,n_j}} = \sum_{i \in \mathbf{M}(\theta_{j,n_j})} \mu_{\theta^i_{j,n_j}}. $$

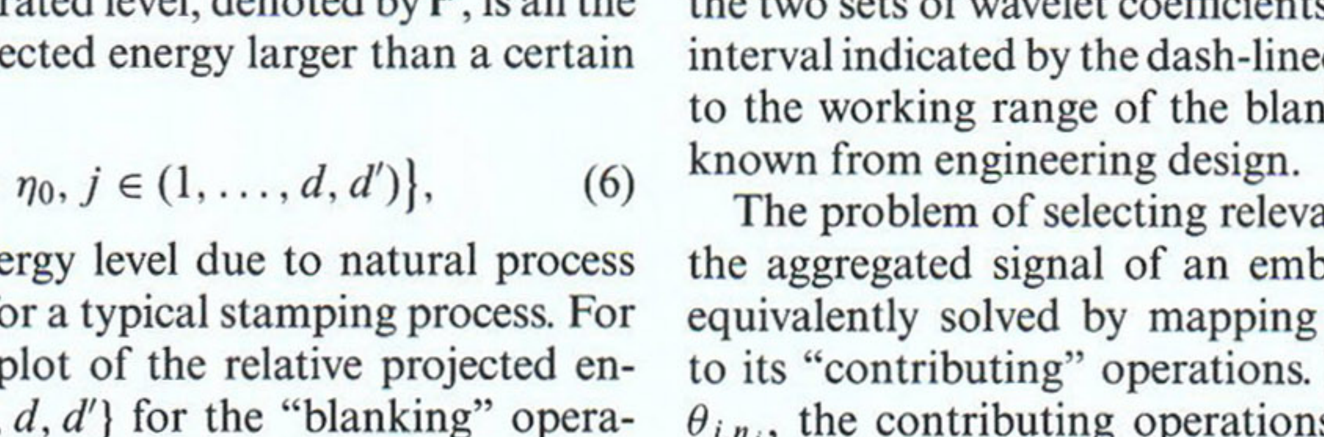

Fig. 6. Comparison of level 2 wavelet coefficients of aggregated signal and those of offline collected blanking signal, indicating that the two sets of wavelet coefficients within the working range of the blanking operation closely match.

Note that $\mathbf{M}(\theta_{j,n_j})$ may be an empty set, indicating that θ_{j,n_j} is not affected by any operations except process noise.

Based on the result of Section 4.1, it is known that because θ_{j,n_j} is a coefficient at level j, it cannot be affected by those operations whose energy does not concentrate on level j, i.e., if $i \in \mathbf{M}(\theta_{j,n_j})$, then $j \in \mathbf{r}^i$. This implies that the search for the contributing operations of θ_{j,n_j} does not need to be conducted throughout all the operations, i.e., $\{0, \ldots, m\}$, but only in a sub-space of $\{0, \ldots, m\}$, which contains all the operations whose energy concentrates on level j. This sub-space is denoted by $\mathbf{\Gamma}(\theta_{j,n_j})$, i.e., $\mathbf{\Gamma}(\theta_{j,n_j}) = \{i | j \in \mathbf{r}^i, i \in (0, \ldots, m)\}$.

To find the contributing operations $\mathbf{M}(\theta_{j,n_j})$ in the sub-space $\mathbf{\Gamma}(\theta_{j,n_j})$, a hypothesis testing procedure is formulated, which integrates engineering design knowledge with statistical testing. Specifically, for each tested set $\mathbf{\Gamma}'(\theta_{j,n_j})$, where $\mathbf{\Gamma}'(\theta_{j,n_j})$ is a subset of $\mathbf{\Gamma}(\theta_{j,n_j})$, i.e., $\mathbf{\Gamma}'(\theta_{j,n_j}) \subseteq \mathbf{\Gamma}(\theta_{j,n_j})$, a hypothesis testing is performed to test the equality of two means: the mean of the wavelet coefficient for the aggregated signal, i.e., $\mu_{\theta_{j,n_j}}$, and the mean of the sum of the wavelet coefficients for the individual operation signals in this tested set, i.e., $\mu_{\theta'_{j,n_j}}$, where

$$\theta'_{j,n_j} = \sum_{i \in \mathbf{\Gamma}'(\theta_{j,n_j})} \theta^i_{j,n_j}.$$

If the testing cannot reject the hypothesis of two means being equal, then $\mathbf{\Gamma}'(\theta_{j,n_j})$ is considered as the set of contributing operations for θ_{j,n_j}, i.e., $\mathbf{M}(\theta_{j,n_j}) = \mathbf{\Gamma}'(\theta_{j,n_j})$.

The hypothesis testing procedure includes the following three steps:

Step 1. Define the test hypotheses, that is

$$\begin{aligned} H_0 &: \mu_{\theta_{j,n_j}} = \mu_{\theta'_{j,n_j}}, \\ H_1 &: \mu_{\theta_{j,n_j}} \neq \mu_{\theta'_{j,n_j}}. \end{aligned} \tag{7}$$

Step 2. Define the test statistic (t statistic), that is

$$t(\theta_{j,n_j}) = \frac{|\hat{\mu}_{\theta_{j,n_j}} - \hat{\mu}_{\theta'_{j,n_j}}|}{\sqrt{(\hat{\sigma}^2_{\theta_{j,n_j}} + \hat{\sigma}^2_{\theta'_{j,n_j}})/K}}, \tag{8}$$

where the sample means and sample variances, i.e., $\hat{\mu}_{\theta_{j,n_j}}$, $\hat{\mu}_{\theta'_{j,n_j}}$, $\hat{\sigma}^2_{\theta_{j,n_j}}$ and $\hat{\sigma}^2_{\theta'_{j,n_j}}$, can be estimated based on samples of the aggregated signal collected online and samples of the individual operation signals collected offline using the experiment in Section 3, and K is the sample size.

Step 3. Determine the decision rule, that is

$$\text{if } t(\theta_{j,n_j}) > t_{\mathrm{T}}, \text{ then } H_0 \text{ is rejected.} \tag{9}$$

In other words, if $t(\theta_{j,n_j}) < t_{\mathrm{T}}$, then we cannot reject the hypothesis that $\mathbf{\Gamma}'(\theta_{j,n_j})$ is the set of contributing operations for θ_{j,n_j}, i.e., $\mathbf{M}(\theta_{j,n_j}) = \mathbf{\Gamma}'(\theta_{j,n_j})$.

From a pure statistical point of view, a statistical threshold for the decision rule can be specified as $t^{\mathrm{S}}_{\mathrm{T}} = t_{\alpha/2,\nu}$ for a given significant level α, i.e., the upper $\alpha/2$ percentage point of a t distribution with ν degrees of freedom, where $\nu = (K+1)(\hat{\sigma}^2_{\theta_{j,n_j}} + \hat{\sigma}^2_{\theta'_{j,n_j}})^2/(\hat{\sigma}^4_{\theta_{j,n_j}} + \hat{\sigma}^4_{\theta'_{j,n_j}}) - 2$ (Montgomery, 2005). Because the individual operation signals are independent, the wavelet coefficients for one signal are independent of those for another signal. Thus

$$\hat{\sigma}^2_{\theta'_{j,n_j}} = \sum_{i \in \mathbf{\Gamma}'(\theta_{j,n_j})} \hat{\sigma}^2_{\theta^i_{j,n_j}}.$$

Based on a reasonable assumption that the variance of a wavelet coefficient of the aggregated signal is equal to the variance of the corresponding wavelet coefficient of each individual operation signal, i.e., $\hat{\sigma}^2_{\theta_{j,n_j}} = \hat{\sigma}^2_{\theta^i_{j,n_j}}$:

$$\hat{\sigma}^2_{\theta'_{j,n_j}} = D(\mathbf{\Gamma}'(\theta_{j,n_j}))\hat{\sigma}^2_{\theta_{j,n_j}}, \tag{10}$$

where $D(\bullet)$ is the cardinality of a set. Thus, $\nu = (K+1)(1 + D(\mathbf{\Gamma}'(\theta_{j,n_j})))^2/(1 + D(\mathbf{\Gamma}'(\theta_{j,n_j}))^2) - 2$.

From an engineering point of view, it is meaningful to conclude only that the two means in the hypothesis testing (7) are different when their difference exceeds natural process variation (i.e., the η_0 defined in Section 4.1). Based on this consideration, an engineering threshold, $t^{\mathrm{E}}_{\mathrm{T}}$, is provided in Proposition 1. (See the online Appendix for the derivation).

Proposition 1. *An engineering threshold for the decision rule in Equation (9) is*

$$t^{\mathrm{E}}_{\mathrm{T}} = \sqrt{\eta_0 K \xi^2/(1+m)}, \tag{11}$$

where η_0 (natural process variation) is available from design and ξ is a "selected" signal-to-noise level.

The selection of an appropriate ξ value is discussed as follows. Because η_0 and m are fixed for a particular stamping process and K is also fixed after the offline experiment is completed, here ξ is the only parameter that can be adjusted to make the engineering threshold statistically meaningful. Specifically, if

$$\xi = \xi_0 \triangleq t_{\alpha/2,\nu}\sqrt{(1+m)/(\eta_0 K)}, \tag{12}$$

then the equality of the engineering and statistical thresholds, i.e., $t^{\mathrm{E}}_{\mathrm{T}} = t^{\mathrm{S}}_{\mathrm{T}}$, is achieved. An interpretation of this result is that we should consider those wavelet coefficients with signal-to-noise ratios no less than ξ_0 to be reliable mapping features, but those with lower signal-to-noise ratios to be unreliable because they contain too much process noise and are thus not statistically meaningful features. In addition, Equation (12) also reveals the relationship between the signal-to-noise ratio ξ_0 and the design parameters η_0, m and K. As an example, Fig. 7 plots the ξ values with respect to different α values for a particular stamping process with $\eta_0 = 10\%$, $m = 4$ and $K = 6$. When $\alpha = 0.1$, $\xi_0 = 5.2$, which means that the search procedure will only map wavelet coefficients with signal-to-noise

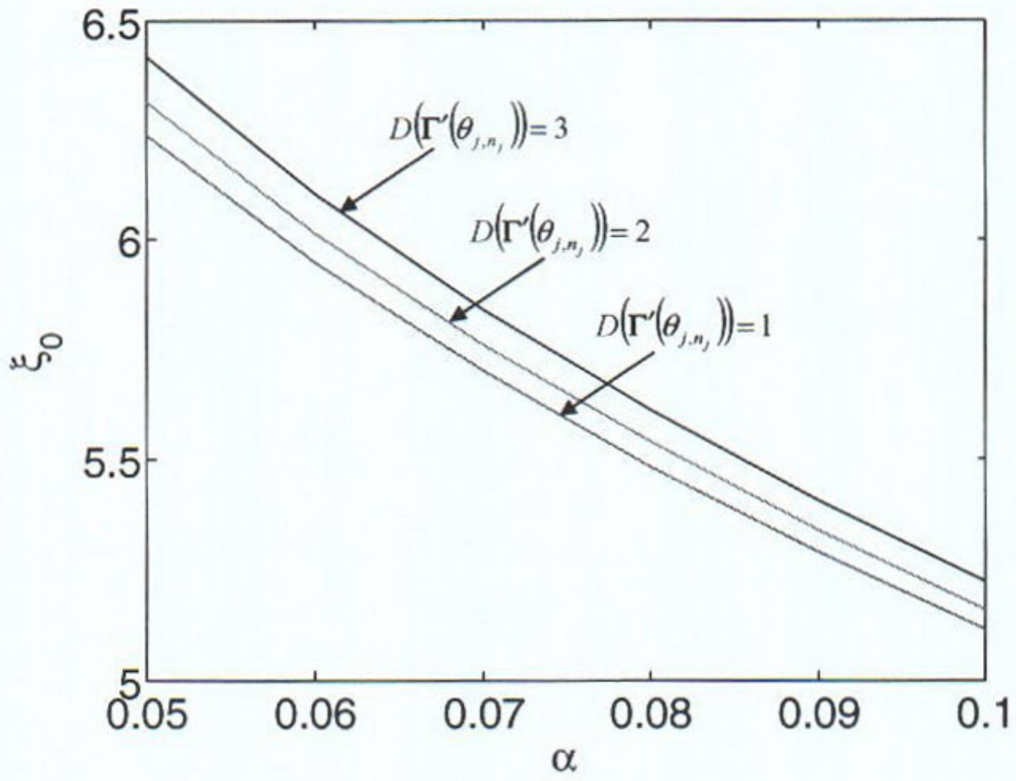

Fig. 7. A decreasing trend between signal-to-noise ratio and significant level for a stamping process with $\eta_0 = 0.1$, $m = 4$ and $K = 6$.

ratios higher than 5.11, 5.16 and 5.22, respectively for $D(\mathbf{\Gamma}'(\theta_{j,n_j})) = 1$, $D(\mathbf{\Gamma}'(\theta_{j,n_j})) = 2$ and $D(\mathbf{\Gamma}'(\theta_{j,n_j})) = 3$ to individual operations, while other wavelet coefficients are considered to correspond to noise at the significant level $\alpha = 0.1$.

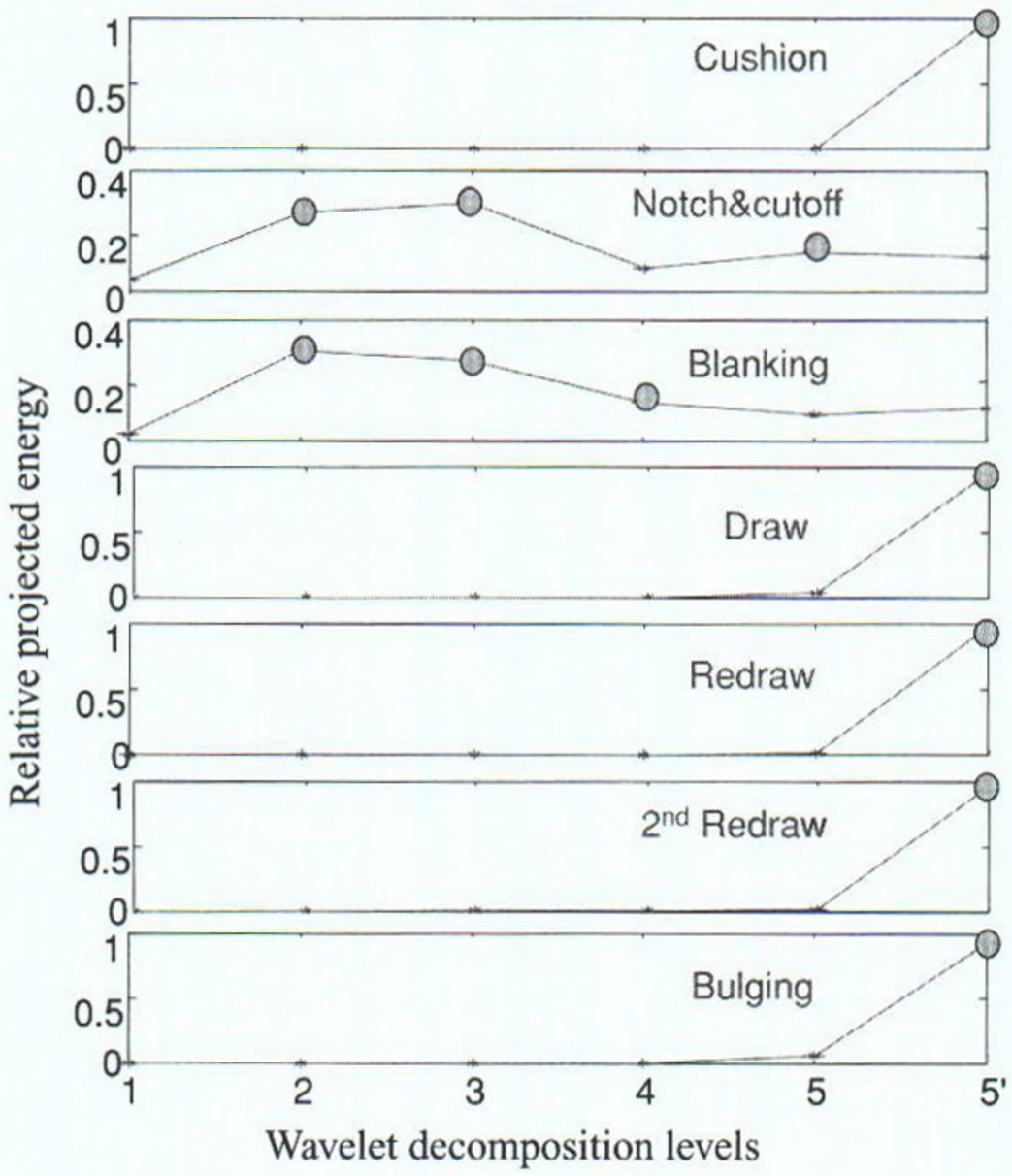

Fig. 8. Wavelet power spectrum analysis for each individual operation, with energy-concentrated levels indicated by solid dots.

5. Case study

A case study is performed on the stamping process in Fig. 1. Six samples (i.e., $K = 6$) for the aggregated tonnage signals were collected online during continuous production, and six samples for each individual operation signal were collected offline based on the experiment in Section 3.

The procedure in Fig. 4 was followed.

First, a DWT (DB4 wavelet basis and maximum decomposition level $d = 5$) was performed on the samples of the aggregated and individual station signals, and the corresponding wavelet coefficients were obtained, respectively.

Second, the wavelet power spectrum analysis was applied to each operation and the relative projected energy at wavelet decomposition levels $j \in \{1, \ldots, 5, 5'\}$ is plotted in Fig. 8. By using $\eta_0 = 10\%$ as the noise energy level, the energy-concentrated wavelet decomposition levels can be identified as $\mathbf{r}^0 = \{5'\}$ for cushion, $\mathbf{r}^1 = \{2, 3, 5\}$ for notch, $\mathbf{r}^2 = \{2, 3, 5\}$ for cutoff, $\mathbf{r}^3 = \{2, 3, 4\}$ for blanking, $\mathbf{r}^4 = \{5'\}$ for draw, $\mathbf{r}^5 = \{5'\}$ for redraw, $\mathbf{r}^6 = \{5'\}$ for second redraw and $\mathbf{r}^7 = \{5'\}$ for bulging. Then, we can identify a sub-space of operations, in which we can search for the contributing operations for each wavelet coefficient of the aggregated signal. This search for a sub-space for each coefficient is identified to be $\mathbf{\Gamma}(\theta_{2,n_2}) = \{1, 2, 3\}$ ($n_2 \in \{1, \ldots, N_2\}$) for all detail coefficients at level 2, $\mathbf{\Gamma}(\theta_{3,n_3}) = \{1, 2, 3\}$ ($n_3 \in \{1, \ldots, N_3\}$) for all detail coefficients at level 3, $\mathbf{\Gamma}(\theta_{4,n_4}) = \{3\}$ ($n_4 \in \{1, \ldots, N_4\}$) for all detail coefficients at level 4, $\mathbf{\Gamma}(\theta_{5,n_5}) = \{1, 2\}$ ($n_5 \in \{1, \ldots, N_5\}$) for all detail coefficients at level 5 and $\mathbf{\Gamma}(\theta_{5',n'_5}) = \{0, 4, 5, 6, 7\}$ ($n'_5 \in \{1, \ldots, N'_5\}$) for all approximation coefficients at level 5.

Third, for each subset of $\mathbf{\Gamma}(\theta_{2,n_2}) = \{1, 2, 3\}$, i.e., $\mathbf{\Gamma}'(\theta_{2,n_2}) \in \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$, the hypothesis testing in Equations (7) to (9) was performed, and the contributing operations of θ_{2,n_2} were found. This creates a mapping between θ_{2,n_2} and the contributing operations, as shown in Table 1. For example, $\theta_{2,19}$, $\theta_{2,20}$ and $\theta_{2,21}$ are mapped to notch; $\theta_{2,26}$, $\theta_{2,27}$ and $\theta_{2,28}$ are mapped to blanking; and $\theta_{2,29}$ is mapped to both blanking and cutoff. Similarly, the mapping between θ_{3,n_3} (also θ_{5',n'_5}) and the individual operations can be found, as given in Table 1.

The results in Table 1 lead to effective strategies in process monitoring and sensor placement in the stamping process. Specifically, because some wavelet coefficients of the aggregated signals are dedicated to a specific operation, including e.g., $\theta_{5',12}$, $\theta_{5',13}$, $\theta_{5',14}$ and $\theta_{5',15}$ being dedicated to cushion; $\theta_{2,19}$, $\theta_{2,20}$, $\theta_{2,21}$ and $\theta_{3,10}$ being dedicated to notch; and $\theta_{2,26}$, $\theta_{2,27}$, $\theta_{2,28}$, $\theta_{3,15}$ and $\theta_{3,17}$ being dedicated to blanking, these coefficients can be used to monitor the performance of the corresponding operation exclusively. This allows for the quality assessment of the individual operations of cushion, notch and blanking, based on the aggregated signals from press sensors, thus avoiding the installation of additional in-die sensors in these operations. Furthermore, for those individual operations lacking dedicated wavelet coefficients such as cutoff, two draws or bulging, Table 1 explicitly shows how those operations are combined to affect

Table 1. Mapping of the wavelet coefficients of the aggregated signal to the operations notch, cutoff, blanking, cushion, 3-draws (draw, redraw and second redraw together) and bulging; the wavelet coefficients not shown here cannot be mapped to any operations, i.e., they correspond to process noise

	Cushion	*Notch*	*Cutoff*	*Blanking*	*3 draws*	*Bulging*
$\theta_{2,19}$		1				
$\theta_{2,20}$		1				
$\theta_{2,21}$		1				
$\theta_{2,26}$				1		
$\theta_{2,27}$				1		
$\theta_{2,28}$				1		
$\theta_{2,29}$			1	1		
$\theta_{3,10}$		1				
$\theta_{3,15}$				1		
$\theta_{3,16}$			1	1		
$\theta_{3,17}$				1		
$\theta_{5',6}$	1				1	
$\theta_{5',7}$	1				1	
$\theta_{5',8}$	1				1	
$\theta_{5',9}$	1				1	1
$\theta_{5',10}$	1				1	1
$\theta_{5',11}$	1					1
$\theta_{5',12}$	1					
$\theta_{5',13}$	1					
$\theta_{5',14}$	1					
$\theta_{5',15}$	1					

the selected mapping wavelet coefficients, specifically, cutoff is combined with blanking, three draws are combined with cushion and bulging, bulging is combined with cushion and three draws. This explicit mapping relationship can still help expedite the search for specific failure stations, even though the dedicated wavelet coefficients cannot be found for those operations. Additionally, it can also help identify the minimum number of additional in-die sensors that are especially needed, such that those operations can be effectively separated so as to enable individual operation performance monitoring.

Verification of the results in Table 1 was conducted by checking whether the time segment involved in computing each wavelet coefficient is within the working range of the operation to which this coefficient is mapped. The verification includes three steps:

Step 1. Identify the working ranges of the stamping operations through engineering design knowledge. Details of the identification procedure can be found in Jin and Shi (2005) and the results are shown in the second row of Table 2.

Step 2. Identify the time segment involved in computing each wavelet coefficient in the first column of Table 2 (same wavelet coefficients as those in Table 1). The identified time segments are shown in Table 2. For example, in computing $\theta_{2,19}$, all points in the aggregated signal within the segment [55 76] are used.

Table 2. Working ranges of individual stamping operations and time segments involved in computing wavelet coefficients

	Working range of each operation (in time units)					
Time segment of wavelet coeff.	*Cushion (entire signal)*	*Notch [64 70]*	*Cutoff [106 113]*	*Blanking [100 113]*	*3 draws [27 113]*	*Bulging [93 187]*
$\theta_{2,19}$		[55 76]				
$\theta_{2,20}$		[59 80]				
$\theta_{2,21}$		[63 84]				
$\theta_{2,26}$				[83 104]		
$\theta_{2,27}$				[87 108]		
$\theta_{2,28}$				[91 112]		
$\theta_{2,29}$			[95 116]	[95 116]		
$\theta_{3,10}$		[31 80]				
$\theta_{3,15}$				[71 120]		
$\theta_{3,16}$			[79 128]	[79 128]		
$\theta_{3,17}$				[87 136]		
$\theta_{5',6}$	[1 192]				[1 192]	
$\theta_{5',7}$	[7 224]				[7 224]	
$\theta_{5',8}$	[39 256]				[39 256]	
$\theta_{5',9}$	[71 288]				[71 288]	[71 288]
$\theta_{5',10}$	[103 300]				[103 300]	[103 300]
$\theta_{5',11}$	[135 300]					[135 300]
$\theta_{5',12}$	[167 300]					
$\theta_{5',13}$	[199 300]					
$\theta_{5',14}$	[231 300]					
$\theta_{5',15}$	[263 300]					

Step 3. Compare the working range of an operation and the time segment of the wavelet coefficient that is mapped to this operation. If the working range and time segment overlap, this wavelet coefficient is a valid mapping feature for this operation. For example, the working range of notch is [64, 70], which overlaps with the time segment of $\theta_{2,19}$. Thus, $\theta_{2,19}$ is a valid mapping feature for blanking. Similarly, it can be verified that all the wavelet coefficients in Table 2 are valid mapping features for the corresponding individual operations.

Furthermore, to evaluate the performance of the proposed method in analyzing signals with different profiles, we conducted two simulations to show the effectiveness of separating aggregated signals in both time and frequency domains. In both studies, the aggregated signal, $\mathbf{S}$, is composed by two individual operation signals, $\mathbf{S}_1$ and $\mathbf{S}_2$, i.e., $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$; $\mathbf{S}_1$ is fixed to be $\mathbf{S}_1 = \sin(t)$, $t \in [11, 80]$, and $\mathbf{S}_1 = 0$, otherwise.

In the first simulation study, $\mathbf{S}_2 = \sin(0.3t)$, $t \geq t_0$ and $\mathbf{S}_2 = 0$, otherwise, where t_0 is given different values, resulting in different profiles of $\mathbf{S}$ in the *time* domain. The procedure in Fig. 4 was followed. The result of this study showed that the less $\mathbf{S}_2$ overlaps with $\mathbf{S}_1$, the more wavelet coefficients can be mapped to $\mathbf{S}_1$. Therefore, aggregated signals consisting of time-localized individual operations can be effectively separated by applying the proposed method.

In the second simulation study, $\mathbf{S}_2 = \sin(ct)$, $t \geq 41$ and $\mathbf{S}_2 = 0$, otherwise, where c is given different values, resulting in different profiles of $\mathbf{S}$ in the *frequency* domain. The procedure in Fig. 4 was followed. The result of this study showed that the more dissimilar the frequency of $\mathbf{S}_2$ is to that of $\mathbf{S}_1$, the more wavelet coefficients can be mapped to $\mathbf{S}_1$. Therefore, the proposed method can be effectively used for mapping aggregated signals into individual operations which generate different localized frequencies in their response signals.

6. Conclusions

This paper developed a method for mapping the features of aggregated signals to the embedded individual operations. A multistage progressive die stamping process was used as an example to demonstrate the developed method. The development of this method included three fundamental components: first, samples of the aggregated signal were collected online from continuous production, together with samples of the individual operation signals which were collected offline through a specially designed physical experiment; second, a DWT was applied to the collected aggregated and individual operation signals and the wavelet coefficients of the signals were obtained; third, a two-step mapping algorithm was performed to map the wavelet coefficients of the aggregated signal to individual operations. Specifically, the first step of this algorithm applied a power spectrum analysis to the wavelet coefficients of the offline-collected individual operation signals and identified the energy-concentrated wavelet decomposition levels for each operation. These energy-concentrated levels correspond to the local frequency range of the operation. In addition to being localized in the frequency domain, each operation is also localized in the time domain. This property is used in the second step of the mapping algorithm, in which the contributing operations for each wavelet coefficient were identified. This two-step algorithm created a mapping between the wavelet coefficients of the aggregated signal and the individual operations. The mapping allows for efficient monitoring and quality assessment of the embedded operations based on the aggregated signals, thereby avoiding installing additional in-die sensors in all operations.

Acknowledgements

This research is partially supported by the NSF Career Award DMI-0549306, NSF grant DMI-0541750 and NSF Engineering Research Center on Reconfigurable Manufacturing Systems at the University of Michigan (NSF grant EEC9529125).

References

Apley, D.W. and Shi, J. (2001) A factor-analysis method for diagnosing variability in multivariate manufacturing processes. *Technometrics*, **43**, 84–95.

Barford, P., Kline, J., Plonka, D. and Ron, A. (2002) A signal analysis of network traffic anomalies, in *Proceedings of ACM SIGCOMM Internet Measurement Workshop*, ACM, New York, NY, 71–82.

Bracewell, R.N. (2000) *The Fourier Transform and its Applications*, McGraw Hill, Boston, MA.

Ceglarek, D. and Shi, J. (1996) Fixture failure diagnosis for autobody assembly using pattern recognition. *Transactions of the ASME, Journal of Engineering for Industry,* **188**, 55–65.

Ding, Y., Zeng, L. and Zhou, S. (2006) Phase I analysis for monitoring nonlinear profiles in manufacturing processes. *Journal of Quality Technology*, **38**(3), 199–216.

Fan, J.Q. and Lin, S.K. (1998) Test of significance when data are curves. *Journal of the American Statistical Association*, **93**(443), 1007–1021.

Gardner, M.M., Mozumder, P.K., Rao, S., Wortman, J.J., Davis, J.C., Heinisch, H.H., Rying, E.A., Lu, J.C., Hornung, B.E. and Gyurcsik, R.S. (1997) Equipment fault detection using spatial signatures. *IEEE Transactions on Components, Packaging, and Manufacturing Technology Part C: Manufacturing*, **20**(4), 295–304.

Guo, H. and Jin, J. (2006) Travel time estimation with correlation analysis of single loop detector data. *Transportation Research Record, Artificial Intelligence and Advanced Computing Applications,* No. 1968, pp. 10–19.

Jin, J. (2004) Individual station monitoring using press tonnage sensors for multiple operation stamping processes. *Transactions of the ASME, Journal of Manufacturing Science and Engineering,* **126**(1), 83–90.

Jin, J. and Shi, J. (1999) Feature-preserving data compression of stamping tonnage information using wavelets. *Technometrics*, **41**(4), 327–339.

Jin, J. and Shi, J. (2001) Automatic feature extraction of waveform signals for in-process diagnostic performance improvement. *Journal of Intelligent Manufacturing,* **12**(3), 257–268.

Jin, J. and Shi, J. (2005) Press tonnage signal decomposition and validation analysis for transfer or progressive die processes. *Transactions of the ASME, Journal of Manufacturing Science and Engineering*, **127**(1), 231–235.

Johnson, R.A. and Wichern, D.W. (2002) *Applied Multivariate Statistical Analysis*, Prentice-Hall, Upper-Saddle River, NJ.

Kang, L. and Albin, S.L. (2000) On-line monitoring when the process yields a linear profile. *Journal of Quality Technology*, **32**(4), 418–426.

Kibarian, J.K. and Strojwas, A.J. (1991) Using spatial information to analyze correlations between test structure data. *IEEE Transactions on Semiconductor Manufacturing,* **4**(3), 219–225.

Kim, K., Mahmoud, M.A. and Woodall, W.H. (2003) On the monitoring of linear profiles. *Journal of Quality Technology,* **35**(3), 317–328.

Lada, E.K., Lu, J.-C. and Wilson, J.R. (2002) A wavelet-based procedure for process fault detection. *IEEE Transactions on Semiconductor Manufacturing*, **15**(1), 79–90.

Li, J., Shi, J. and Chang, T.S. (2007) On-line seam detection in rolling processes using snake projection and discrete wavelet transform. *Transactions of the ASME, Journal of Manufacturing Science and Engineering,* **129**(5), 926–933.

Mahmoud, M.A., Hawkins, D.M., Parker, P.A. and Woodall, W.H. (2007) A change point method for linear profile data. *Quality and Reliability Engineering International*, **23**(2), 247–268.

Mahmoud, M.A. and Woodall, W.H. (2004) Phase I analysis of linear profiles with calibration applications. *Technometrics*, **46**(4), 380–391.

Mallat, S.G. (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, **11**, 674–693.

Montgomery, D.C. (2005) *Introduction to Statistical Quality Control*, Wiley, New York, NY.

Rao, C.R. and Kleffe, J. (1988) *Estimation of Variance Components and Applications*, North-Holland, Amsterdam, The Netherlands.

Stover, F.S. and Brill, R.V. (1998) Statistical quality control applied to ion chromatography calibrations. *Journal of Chromatography A*, **804**(1–2), 37–43.

Walker, E. and Wright, S.P. (2002) Comparing curves using additive models. *Journal of Quality Technology*, **34**(1), 118–129.

Winistorfer, P.M., Young, T.M. and Walker, E. (1996) Modeling and comparing vertical density profiles. *Wood and Fiber Science*, **28**(1), 133–141.

Woodall, W.H., Gupta, S., Spitzner, D.J. and Montgomery, D.C. (2004) Using control charts to monitor process and product quality profiles. *Journal of Quality Technology*, **36**(3), 309–320.

Young, T.M., Winistorfer, P.M. and Wang, S. (1999) Multivariate control charts of MDF and OSB vertical density profile attributes. *Forest Products Journal*, **49**(5), 79–86.

Zhou, S., Ding, Y., Chen, Y. and Shi, J. (2003) Diagnosability study of multistage manufacturing processes based on linear mixed-effects models. *Technometrics,* **45**, 312–325.

Biographies

Jionghua (Judy) Jin is an Associate Professor in the Department of Industrial and Operations Engineering at the University of Michigan. She received her Ph.D. degree from the University of Michigan in 1999. Her recent research focuses on data fusion for complex system improvement with the goal of developing novel methodologies for variation reduction, condition monitoring and fault diagnosis, process control, knowledge discovery and decision making. Her research emphasizes a multidisciplinary approach by integrating applied statistics, signal processing, reliability engineering, system control and decision-making theory. She has received a number of awards including the NSF CAREER Award in 2002 and the PECASE Award in 2004 and five Best Paper Awards during 2005–2008. She is a member of ASME, ASQC, IEEE, IIE, INFORMS and SME.

Jing Li is an Assistant Professor in the Department of Industrial Engineering at Arizona State University. She received her Ph.D. from the University of Michigan in 2007. Her research interests include applied statistics, data mining and causal modeling and inference for process control. Her recent research focuses on modeling and analyzing massive high-dimensional datasets in complex systems in order to improve the quality of products and processes. Her work has been applied to manufacturing and public health problems. She is a member of IIE, INFORMS and ASQ.