HACCP with multivariate process monitoring and fault diagnosis techniques: application to a food pasteurization process

Figen (Kosebalaban) Tokatli [a,1], Ali Cinar [a,*], Joseph E. Schlesser [b]

[a] Department Chemical and Environmental Engineering, Illinois Institute of Technology, Chicago, IL 60616, USA
[b] National Center for Food Safety and Technology, US Food and Drug Administration, Summit-Argo, IL 60501, USA

Abstract

Multivariate statistical process monitoring (SPM), and fault detection and diagnosis (FDD) methods are developed to monitor the critical control points (CCPs) in a continuous food pasteurization process. Multivariate SPM techniques effectively use information from all process variables to detect abnormal process behavior. Fault diagnosis techniques isolate the source cause of the deviation in process variable(s). The methods developed are illustrated by implementing them to monitor the critical control points and diagnose causes of abnormal operation of a high temperature short time (HTST) pasteurization pilot plant. The detection power of multivariate SPM and FDD techniques over univariate SPM techniques is shown and their integrated use to ensure the product safety and quality in food processes is demonstrated.
© 2004 Published by Elsevier Ltd.

Keywords: Multivariate statistical process monitoring; Fault diagnosis; HACCP

1. Introduction

Preservation of the nutritional value and appeal of food products often requires minimal treatment. However, low acid food products are usually subjected to long heating times at elevated temperatures to assure food safety. Powerful and reliable safety monitoring techniques could significantly reduce the processing times while providing the necessary safety margin. When unsafe food products are detected, they should be diverted and production must be halted until the conditions causing unsafe products are diagnosed and eliminated. Consequently, fault diagnosis methods that can quickly identify the source causes of process operation that yield unsafe products will reduce the process down time and productivity loss.

In many food processing operations product safety is controlled by checking only the end product by microbiological and chemical methods. A major drawback associated with this approach is time delay. Collecting and examining the samples to determine the safety of the product takes too much time. It can be a high-cost solution if any contamination is reported after the production is completed. Furthermore, the recall of the defective product and the collection from retail outlets add significant cost. When the inadequacies of traditional food safety control based on end product control have been noticed, more effective ways to control the safety of food processing lines have been sought. The systematic and scientific approach called hazard analysis and critical control points (HACCP) was first used in 1960s and it has been developed further in the 30 year period (Khandke & Mayes, 1998; McAnelly, 1994; Motarjemi, Kaferstein, Moy, Miyogawa, & Miyagishima, 1996; Savage, 1995). HACCP is the way of identifying and controlling hazards during production to ensure that food is safe when it is consumed and it does not present an unacceptable risk to health. Instead of checking only the properties of the end product, the critical control points (CCP) in the process are monitored continuously to prevent a possible major hazard in advance. Therefore reliable methods are needed for on-line monitoring of CCPs.

In food processing plants where HACCP plans are applied, the critical limits on specific measured variables

---

* Corresponding author. Fax: +1-312-567-8874.
E-mail address: cinar@iit.edu (A. Cinar).
[1] Current address: Izmir Institute of Technology, Food Engineering Department, Izmir 35430, Turkey.

(the CCP) are used to ensure the safety of the product. Critical control limits are absolute limits. Any measurement outside the critical limit indicates insufficient treatment. In addition to monitoring these hard limits, statistical process monitoring (SPM) charts can be set up on the safe side of the critical control limit and monitored for detecting trends that may eventually cause the process to violate a critical control limit (Fig. 1). One of the control limits of the SPM chart is close to the critical safety limit. A change in critical control variable values that exceeds the SPM chart limit towards the critical limit warns the plant personnel. This early warning provides the opportunity for a timely intervention by plant personnel for corrective action that may prevent the occurrence of noncompliance. Achieving compliance to the critical limit within an acceptable operating range is important for quality maintenance and profitable operation. Keeping the process within the SPM chart interval prevents over-processing which deteriorates the physical properties of the end product and excessive use of energy.

SPM involves the use of statistical techniques to monitor the variability of a process. SPM is often implemented by using monitoring charts that display the intervals associated with control limits and monitor the process on-line. The purpose of SPM is the detection of the out-of-control status and its time of occurrence. While separate charts for each important variable have been developed in many practices (the so-called univariate SPM), process monitoring by integrating information from all important process variables has better performance. This *multivariate* SPM framework is gaining rapid acceptance in many industries and food processing industry can benefit from the advantages of multivariate SPM as well.
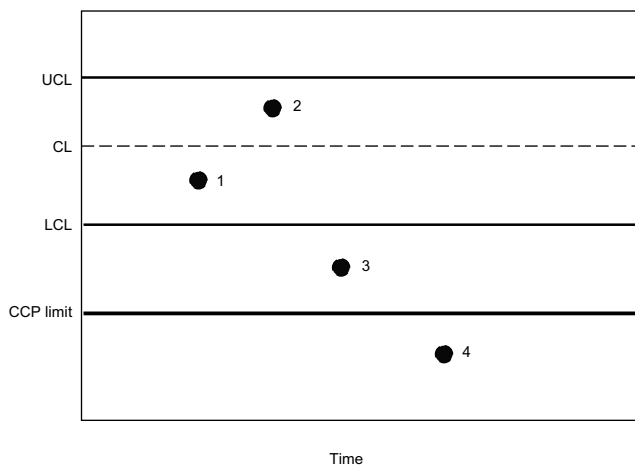

Fig. 1. The CCP safety limit (bold solid line) and SPM upper and lower control chart limits (UCL, LCL) (solid lines) for a CCP. Points 1 and 2 indicate in-control status, 3 is out-of-control and a warning will be issued to plant personnel, and 4 is outside the safety limit, necessitating intervention to divert the product.

Fault diagnosis is a complementary task to SPM that focuses on the identification of process and equipment faults that are the source cause of the abnormality observed. While the SPM chart detects abnormal process operation, fault diagnosis identifies the process variable(s) that caused the system to deviate from its normal operating conditions and the source cause that created the significant deviation in these variables. This information can help the operator for determining the appropriate course of action quickly and making the necessary intervention or repair.

One contribution of this paper is the introduction of SPM within the framework of HACCP. Another contribution is the illustration of the integration of multivariate SPM and fault diagnosis to generate a powerful tool for rapid determination of the cause(s) of abnormal process operation.

The structure of the paper is as follows. Various HACCP, SPM and FDD applications in food processing industry are summarized in Section 2. Multivariate SPM and FDD techniques are presented in Section 3. The HTST pasteurization system and its data acquisition unit are described in Section 4. The performance of the multivariate SPM and fault diagnosis techniques are illustrated by implementing SPM and FD to the HTST pasteurization system and the results are discussed in Section 5.

2. HACCP, SPM and FDD practice in food processing

The adoption of HACCP programs by the food industry is increasing. Today, HACCP plans are being used in dairy, fish, meat, bakery and beverage industries. In the USA, Food and Drug Administration (FDA) and the Food Safety Inspection Service (FSIS) of USDA have been preparing regulations regarding the use of HACCP plans in food and meat processing plants (Bakka, 1998; Bernard, 1998). However, the literature indicates that the use of powerful multivariate SPM and FDD techniques and their integration with HACCP are very limited at the present time in food industry.

In various studies that are listed in the following sections, mostly *univariate* SPM charts such as Shewhart charts (also called x-bar and range charts) are used to monitor CCPs and these charts are combined with a FDD tool. Besides monitoring CCPs, there are examples of using univariate SPM charts for quality characteristics of food products. It is apparent that the multivariate SPM and FDD methods have not yet found many application areas in the food industry.

2.1. Univariate SPM in HACCP and in quality control

Monitoring CCPs in the process individually by univariate SPM charts is suggested in HACCP (Hub-

bard, 1990; McAnelly, 1994; Ozilgen, 1998b). However, it is not clear in the literature that this has been practiced extensively. Use of statistical univariate cumulative sum (CUSUM) and Shewhart charts to monitor a CCP in a milk-filling machine has been reported. The purpose of the study was to establish a trend analysis of CCP data by treating data with appropriate statistical tools in order to obtain advanced warning on the status of the process, not just a pass/fail classification (Hayes, Scallon, & Wong, 1997).

Univariate SPM charts are being used for quality control in food industry. In sugar production from beet, Shewhart charts have been developed for the control of some product parameters within the manufacturing process itself after a training period (Sanigar, 1990). The company that implemented this effort is considered as one of the pioneers in augmenting, SPM with the quality management system. It is reported that the product quality has improved and the annual savings increased significantly after SPM applications in the plant. Another example is the use of univariate moving mean and range SPM charts in a confectionery industry to overcome the product inconsistency in texture and size (Bidder, 1990). Statistical monitoring charts were used on-line in peanut butter production at a Procter and Gamble Plant (Miller & Balch, 1991). Pareto charts were produced to decrease the unscheduled down-time since product quality was affected negatively after each start-up. Process capability analysis was used to monitor and control the peanut butter color by roasting the peanuts within specifications and with less variability. Also, mean and range charts were used to regulate the salt addition into the grinder. It was concluded that the use of on-line SPM techniques led to significant process improvement. In another study, the mean and range charts were used in a brewery industry for quality control (Ozilgen, 1998a). Charts were constructed for characteristic quality factors such as total acidity, pH, alcohol content and carbon dioxide after transforming non-normally distributed data to normally distributed data. Another quality control application was reported on hazelnut sorting and cracking unit (Ozdemir & Ozilgen, 1997). Control charts based on percent defective product (p charts) (Hubbard, 1990; Montgomery, 1991) were constructed to monitor and control the performance of two different sizer-cracker units. Statistical quality control tools for food processing and application examples are discussed in two recent books (Hubbard, 1990; Ozilgen, 1998b). However, the contents of these books are limited to univariate statistical charts.

2.2. Multivariate SPM in HACCP and quality control

The importance of multivariate statistical analysis and control in food industry has been pointed out in the literature. However, the application of these methods to a process line was not reported. In an early publication (Buco, 1990), it is stated that quality control in food industry requires more than acceptance sampling and univariate control charts. When there are a number of quality characteristics, the multivariate Hotelling's T^2 chart is suggested to monitor the process variables jointly since multivariate charting techniques are more powerful for detecting simultaneous changes. More comprehensive research on multivariate SPM is presented in Negiz, Ramanauskas, Cinar, Schlesser, and Armstrong (1998a, 1998b, 1998c) where empirical model development and multivariate SPM techniques for a HTST pasteurization process are discussed. An SPM method based on the state variables of the canonical variate state space model of the process is used. A T^2 chart of state variables is used to detect the abnormalities in the process variables. Negiz and Cinar (1998) have also reported statistical monitoring charts for sensor fault detection that indicate whether there is bias, drift or increased noise in the measurements of individual sensors. T^2 and SPE charts of state variables and contribution plots of process variables are interfaced with a knowledge-based system (KBS) to integrate and automate process monitoring and fault diagnosis (Norvilas, Negiz, DeCicco, & Cinar, 2000). This framework was used to monitor a sausage cooking unit simulated in a real-time KBS environment (Norvilas & Cinar, 1997). Multivariate statistical monitoring of a sausage cooking line based on plant data is presented in DeCicco, Martino, Cinar, Verdoorn, and Balasubramaniam (1999). The external sausage temperature was measured with infrared thermometer and a linear relationship was developed between external sausage temperature and internal sausage temperature which is the CCP for cooked sausage processing. Temperature readings were taken across the conveyor belt at each sampling time to represent the temperature distribution. Multivariate SPM tools for external sausage temperature across the conveyor belt are based on the Hotelling's T^2 chart. Univariate charts are also used to provide additional information for diagnosis.

2.3. Automated FDD in food industry

A number of FDD applications in food industry based on the use of KBS have been reported. Alanso, Acosta, Prada, and Mira (1994) describe an expert system for on-line FDD in a beet sugar plant. Monitored variables of the process are checked against an interval. When monitoring indicates a problem with a variable, diagnosis task is initiated. Similar applications have been done in the bottling line in a brewery company (Troupis, Manesis, Koussoulas, & Chronopoulos, 1995) and in the juice purification unit in a cane sugar production (Pokkunuri, 1994). In both applications, univariate SPM charts are used to detect a problem

and the KBS is used to handle the fault diagnosis activity.

3. Multivariate SPM and FDD techniques

Industrial processes have many variables. SPM and FDD techniques that make use of the information contained in all variables have resulted in faster and better monitoring and diagnosis. Multivariate SPM and FDD techniques are presented in this section.

3.1. Multivariate SPM techniques

When more than one process or quality variable is measured at a time, the univariate charts of individual variables are not preferred because they ignore the correlation in data and the additional information provided by the other variables. Interactions among variables do not allow individual variables to behave independently. The collective information from several variables often indicates a burgeoning trend towards deviation earlier than any individual variable.

The Hotelling's T^2 chart is a widely accepted multivariate SPM tool. It measures the cumulative deviation of variables in a data set from their mean values based on the concept of statistical distance. T_k^2 value at time k is given as

$$T_k^2 = (\mathbf{y}_k - \boldsymbol{\mu})^{\mathrm{T}} \mathbf{S}^{-1} (\mathbf{y}_k - \boldsymbol{\mu}), \tag{1}$$

where $\mathbf{y}_k$ is the observation vector at time k, $\boldsymbol{\mu}$ is the mean vector or target value and $\mathbf{S}$ is the in-control covariance matrix of the observations $\mathbf{y}$. The superscript T denotes the transpose of a matrix and $\mathbf{S}^{-1}$ is the inverse of matrix $\mathbf{S}$.

When variables are highly cross-correlated and collinear, additional mathematical steps must be considered. The trajectories of collinear variables show strong similarity and cause numerical difficulties in computing the inverse of $\mathbf{S}$. These numerical problems in computing $\mathbf{S}^{-1}$ will yield a T^2 chart that gives too many false alarms. Hence, the T^2 chart of process variables may not be a suitable choice for processes with highly cross-correlated and collinear variables. Principal component analysis (PCA) can be used to remedy this problem. In addition, data may have strong autocorrelation. In recent works, SPM based on state variables of a dynamic stochastic model of the process were developed from process data collected under normal operating conditions to deal with high autocorrelation and cross-correlation in process data (Negiz & Cinar, 1997; Norvilas et al., 2000). SPM of processes generating highly autocorrelated and cross-correlated data can be implemented by using the T^2 chart of canonical variate state space variables and the squared prediction error (SPE) chart.

The multivariate T^2 chart of state variables is used in detecting the abnormal process behavior. The T^2 statistic for state variables at each time k is distributed as F distribution and is computed as

$$T_k^2 = \mathbf{x}_k^{\mathrm{T}} \mathbf{S}_{\mathrm{st}}^{-1} \mathbf{x}_k \sim \frac{n(N^2 - 1)}{N(N - n)} F_{\alpha, n, N-n}, \tag{2}$$

where $\mathbf{x}$ is the state variables vector, $\mathbf{S}_{\mathrm{st}}$ is the covariance matrix of state variables, N is the number of observations, n is the number of state variables. F denotes F distribution and α is the confidence limit.

If the dynamic model cannot explain the data collected at measurement time k, the statistics of the new observations will move away from the space defined by the in-control historical data and dynamic process model. In other words, the existing process model cannot explain the particular observation vector adequately. The residual vector $\mathbf{e}_k$ computed by using process data $\mathbf{y}_k$ and its estimates from the state variables model $\hat{\mathbf{y}}_k$, shows the goodness of the fit of the model to data

$$\mathbf{e}_k = \mathbf{y}_k - \hat{\mathbf{y}}_k. \tag{3}$$

The residuals are monitored by the normalized SPE chart (SPE$_N$). At time k, the SPE$_{N_k}$ value and its distribution are

$$\begin{aligned} \mathrm{SPE}_{N_k} &= (\mathbf{e}_k - \bar{\mathbf{e}})^{\mathrm{T}} \mathbf{S}_e^{-1} (\mathbf{e}_k - \bar{\mathbf{e}}), \\ \mathrm{SPE}_N &\sim \frac{p(N^2 - 1)}{N(N - p)} F_{\alpha, p, N-p}, \end{aligned} \tag{4}$$

$\mathbf{S}_e$ and $\bar{\mathbf{e}}$ are the covariance matrix and the mean vector of residual matrix respectively, which are determined from the in-control data. The in-control residual mean vector $\bar{\mathbf{e}}$ is almost zero and in-control residual covariance matrix $\mathbf{S}_e$ is diagonal (uncorrelated residuals).

The multivariate SPM techniques monitor two statistics (T_k^2 and SPE$_{N_k}$) at each measurement time k by using all the process variables. They signal that the process is out-of-control when either statistic exceeds its upper limit. The multivariate SPM charts do not provide information on which variables are out-of-control. After detecting the out-of-control status, fault diagnosis methods are used to identify the variable(s) that signals the deviation from normal behavior.

3.2. Multivariate FDD techniques

FDD techniques isolate and diagnose the source of abnormality detected by the multivariate monitoring charts. There are different ways of implementing FDD. The methods may be collected under two main titles: model-free and model-based methods. The model-free methods do not require a mathematical model of the system. FDD with model-free methods can be performed through simple limit checking of plant mea-

surements, comparison of the same physical quantity from different sensors (physical redundancy) or spectrum analysis of plant measurements (Gertler, 1998). The model-based methods are also named as analytical redundancy or quantitative FDD. These methods are based on a mathematical model of the system that provides estimates of process data. The idea is, to compare the measurements with analytically computed values and infer from significant differences the variables that have affected the statistics that indicated the deviation. Either by using process knowledge or additional diagnosis methods, these variables are, then linked to specific process equipment whose malfunction would cause the effects observed on these variables.

The fault diagnosis technique used in this study is called the *parity space technique*, which is a model-based FDD method (Peng, Youssouf, Arte, & Kinnaert, 1997). The technique utilizes parity residuals that are computed by using 'balance equations' based on input and output data of the process and parity relations. The first step is the development of the dynamic process model between process inputs and outputs. In the second step, the parity relations are developed and the procedures are formulated to monitor parity residuals that indicate whether the parity relations are satisfied or not. In this work, parity residuals are tested with the generalized likelihood ratio (GLR) test (Johnson & Wichern, 1998; Willsky & Jones, 1976).

Parity residuals are generated from input and output measurements by using a dynamic, discrete time state space model of the system:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{J}\mathbf{d}_k + \mathbf{E}\mathbf{f}_k,$$
$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{P}\mathbf{d}_k + \mathbf{F}\mathbf{f}_k, \tag{5}$$

where $\mathbf{x}_k$ denotes the state variables vector, $\mathbf{u}_k$ is the actuator command vector, $\mathbf{y}_k$ is the output variables vector, $\mathbf{d}_k$ is the white noise vector, and $\mathbf{f}_k$ is the fault modes vector. $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, $\mathbf{E}$, $\mathbf{F}$, $\mathbf{J}$, and $\mathbf{P}$ are constant system matrices.

The fault modes in the parity space technique are classified as sensor faults and actuator faults. Different parity relations for sensor faults and actuator faults are defined by using the dynamic process model (Eq. (5)). Parity residuals for sensor faults are

$$\mathbf{r}_{i_k} = \mathbf{H}(z)\mathbf{y}_{i_k} + \mathbf{K}(z)\mathbf{u}_k, \tag{6}$$

where $\mathbf{r}_{i_k}$ denotes the parity residual vector for the ith sensor at time k, $\mathbf{H}(z)$ is the transfer function between output measurements and residuals, and $\mathbf{K}(z)$ is the transfer function between input measurements and residuals. The argument (z) indicates that the transfer functions are in z domain.

Parity residuals for actuator faults are generated similar to Eq. (6):

$$\mathbf{r}_{i_k} = \mathbf{P}(z)\mathbf{y}_k + \mathbf{R}(z)\mathbf{u}_{i_k}, \tag{7}$$

where $\mathbf{r}_{i_k}$ denotes the parity residual vector for the ith actuator at time k, $\mathbf{P}(z)$ is the transfer function between output measurements and residuals, and $\mathbf{R}(z)$ is the transfer function between input measurements and residuals.

For each process variable, a parity relation is generated and the parity residuals are tested with the GLR test to detect the change in that particular variable (Peng et al., 1997; Willsky & Jones, 1976). Parity residuals are noise series in the absence of fault. When a fault occurs in a particular variable, its parity residual gets a significantly higher value than the noise series. The GLR test is needed to differentiate a fault mode from the noise series more effectively.

The appeal of the parity space approach is its ability to diagnose each fault that has its parity relation. However, faults that do not have parity relations cannot be identified. Often a subset of all possible faults is selected based on the importance of the fault and its parity relations are defined to diagnose them. An alternative is to use the contribution plots that indicate the process variables that have contributed to the out-of-control signal (by inflating T^2 or SPE statistics). Then, process knowledge is used to identify process equipment that will affect these variables when they malfunction in order to deduce the source causes. This approach and its implementation to the HTST pasteurization process are discussed in detail in Kosebalaban (2000). The reasoning (inferencing) to relate process variables with significant variation to process equipment can be automated by using KBSs.

4. HTST pasteurization system

The multivariate process monitoring and diagnosis techniques are illustrated by implementing them on the operation of a HTST pasteurization pilot plant. The HTST pasteurization system is located at the National Center for Food Safety and Technology (NCFST) of Illinois Institute of Technology. A process flow and instrument diagram is given in Fig. 2. The pilot pasteurization plant consists of a plate heat exchanger, a centrifugal pump, a flow diversion valve (FDV), a boiler and a homogenizer. The heat exchanger is a multipass, plate heat exchanger by APV Company. The homogenizer (APV) has an adjustable flowrate between 56.78 and 189.27 l/h (15–50 gal/h). There are two regulatory valves, the steam injection valve (steam valve) to the boiler and the hot water flow valve in preheater section (preheater valve).

The product is heated by hot water. The hot water is heated by direct steam injection in the hot water heater. Three PID control loops are used to control product
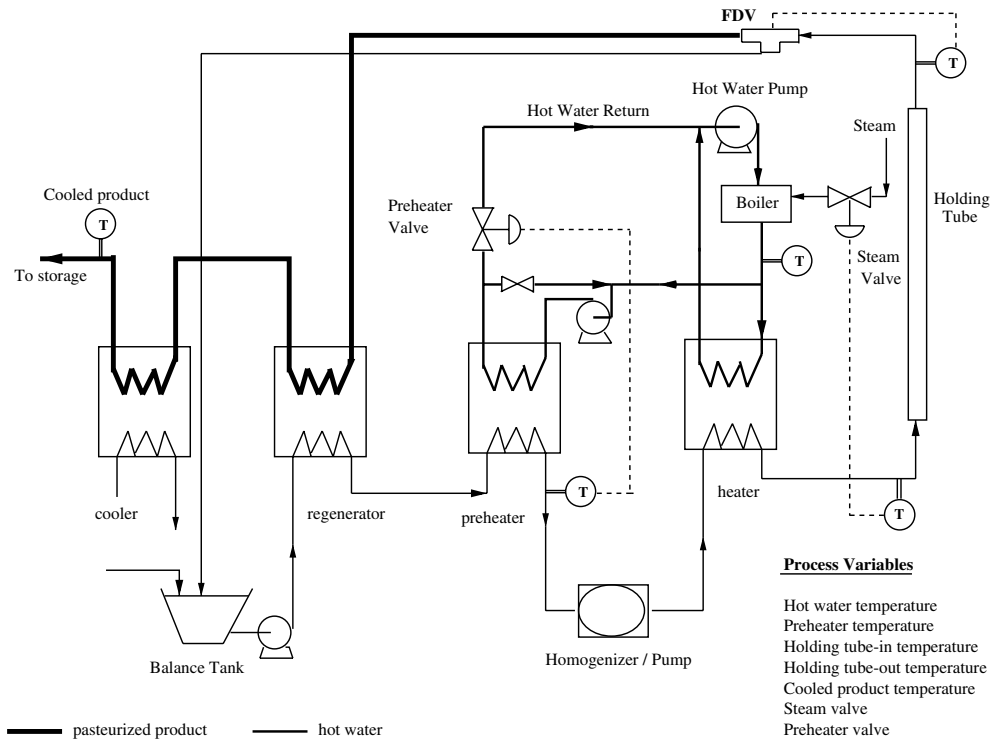
Fig. 2. NCFST pilot pasteurization plant.

temperature. The first control loop regulates the raw product temperature leaving the preheater. The second loop controls product temperature entering the holding tube. The last loop controls the temperature of the pasteurized product leaving the cooler. The raw product temperature at the exit of the preheater is controlled by manipulating the flow of hot water through the preheat heater exchanger. The product temperature at the holding tube inlet is controlled by manipulating the steam flow rate into the hot water heater. The product temperature at the exit of the cooler is controlled by manipulating the flow rate of cold water flowing through the cooler heat exchanger. This control loop is not indicated in Fig. 2 or used in our studies since it has no effect on the pasteurization process. The flow diversion valve is controlled by pasteurized milk temperature at the holding tube exit. The measured variables are hot water, holding tube inlet, holding tube outlet and preheater exit temperatures, and the steam valve and the preheater valve signals. The controlled variables are the holding tube-inlet temperature and the preheater temperature. The manipulated variables are the steam flow rate by steam control valve and the preheater hot water flow rate by preheater control valve.

The HTST system was connected to a computer system equipped with a data acquisition system (DAS) wired to HTST system sensors and actuators. The real-time data acquisition computer system consists of a Hewlett-Packard HP 75000 System with a HP 75000 card cage and a Compaq Presario 7222 personal computer. The measurement card on HP 75000 includes a 5.5 digit multimeter, two 16-channel thermocouple relay multiplexers, a quad 8-bit digital input–output card and a 4-channel digital to analog (D/A) converter. The software used for data acquisition, display, storage and process control is HPVEE 3.12. The control signals which are signals to pneumatic steam valve in main heating loop and preheater valve are generated by proportional–integral–derivative (PID) controllers developed in HPVEE software. Their values are sent to the control valves of the HTST system after the digital signals are first converted to 4–20 mA current signals by D/A converters, then converted to 0.204–1.020 atm (3–15 psig) pneumatic signals by current-to-pneumatic (I/P) transducers.

5. Case study: Sensor fault detection and diagnosis in a continuous food pasteurization process

Detection and diagnosis of sensor faults, actuator faults, or a combination of such faults were implemented and investigated by Kosebalaban (2000). In this study, only sensor faults will be discussed to illustrate the performance of the multivariate SPM and FD techniques presented and to compare them with univariate SPM techniques.

Sensor Faults: Sensor faults are introduced to the system by modifying sensor data values received by the process control computer. To generate a sensor fault for

a specific time period, a real number is added to the actual sensor reading, which is transmitted to the computer from the process sensors. Instead of the actual reading of the sensor, the modified sensor 'reading' is sent to the PID controllers. The controllers computed the manipulated variable settings based on the faulty sensor reading. Therefore, the system received a false action command from the controllers even though there was no need for any adjustments in the flow rates of steam and hot water, and the fault 'implemented' on the sensors was propagated through the system.

The magnitudes of faults implemented to the holding tube-inlet temperature sensor, were +0.39 and +0.83 °C (+0.7 and +1.5 °F). The fault duration was between 2 and 8 s. The details of each fault (times, magnitudes and durations) are given in Table 1. Only positive deviations are considered as faults because they represent the more dangerous type of deviation, the temperature sensor reporting values higher than the actual product temperature. This may indicate a safe temperature for pasteurization when in reality the temperature is lower.

5.1. Discussion of detection and diagnosis of sensor faults

The performance of monitoring charts and GLR tests of parity relations are given in Table 2. In this table, abbreviations HW, HT-in, Pht.T., HT-out, St.V. and Pht.V. stand for the hot water, holding tube-inlet, preheater product and holding tube-outlet temperatures and steam valve and preheater valve signals, respec-

tively. The Hotelling's T^2 chart and SPE_N chart detected all faults one sampling time after the faults were initiated in the system (Fig. 3). The instants when faults were implemented are shown by downward arrows in the figures. Shewhart charts and the GLR tests of parity residuals for each variable are given in Figs. 4–9. In these figures, the upper graph shows the observations of the process variables and the univariate SPM chart control limits (2 standard deviation in dashed lines, 3 standard deviation in solid lines). The second graph in each figure (Figs. 4–9) shows the result of FDD with its GLR control limit. The parity relation of the faulty sensor detected the fault at the same time as the multivariate monitoring charts (Fig. 5). Therefore, the alarms in the multivariate SPM charts are diagnosed as the faults in the sensors whose parity relation issues alarms at the same time instant. The univariate Shewhart charts did not alarm many faults either because the fault was of small magnitude (even if it persisted for some time) or the fault was within the in-control variation of the variable and disappeared after certain time. Faults with magnitude of +0.39 °C (+0.7 °F) cannot be detected in any Shewhart chart. Both the multivariate statistics (T^2 and SPE) and GLR tests are much better than the univariate charts in detecting sensor faults with small bias shifts.

The signals of the GLR tests for the actuators indicated the faults a number of sampling times after the GLR signals of the holding tube-inlet temperature measurements. This was an indication that the failure originates from the holding tube-inlet temperature sensor. In case of faults with a magnitude of +0.7 °F in the holding tube-inlet temperature sensor, the GLR chart of the steam valve signal did not give out-of-control alarms because of the small fault magnitude (Fig. 8). The GLR chart of the hot water temperature sensor did not alarm these small magnitude faults either (Fig. 4) since the steam valve did not cause big disturbances in the process.

Sensor faults with positive deviations were used in fault tests of the holding tube-inlet temperature sensor to illustrate the synergy between monitoring CCP limit

Table 1
Holding tube-inlet temperature sensor fault: times and magnitudes of faults

Fault	Fault time (s)	Fault magnitude (F)	Duration (s)
1	150	+0.7	2
2	302	+1.5	2
3	454	+0.7	4
4	608	+1.5	4
5	762	+0.7	8
6	920	+1.5	8

Table 2
Holding tube-inlet temperature sensor fault: results of SPM charts and GLR test of parity relations in terms of the number of measurements before the detection

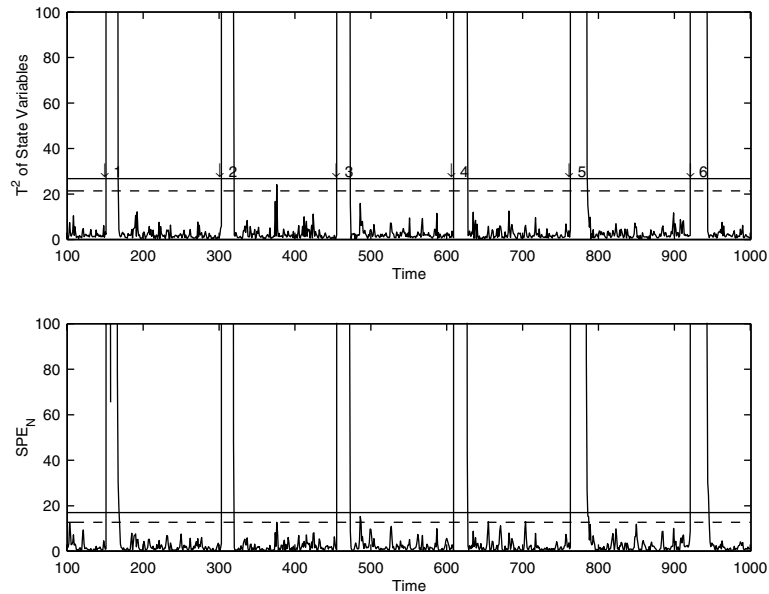| Fault (time)	SPM Results		GLR Results					
	T^2	SPE_N	HW	HT-in	Pht.T.	HT-out	St.V.	Pht.V.
1 (150)	1	1	NA	1	NA	NA	NA	13
2 (302)	1	1	NA	1	NA	77	27	13
3 (454)	1	1	NA	1	NA	NA	NA	NA
4 (608)	1	1	8	1	NA	76	22	12
5 (762)	1	1	NA	1	NA	26	36	23
6 (920)	1	1	16	1	NA	NA	8	13

NA: no alarm generated.

Fig. 3. Holding tube-inlet temperature sensor fault: T^2 of state variables and SPE$_N$ chart with 99% (—) and 95% (- - -) confidence limits.
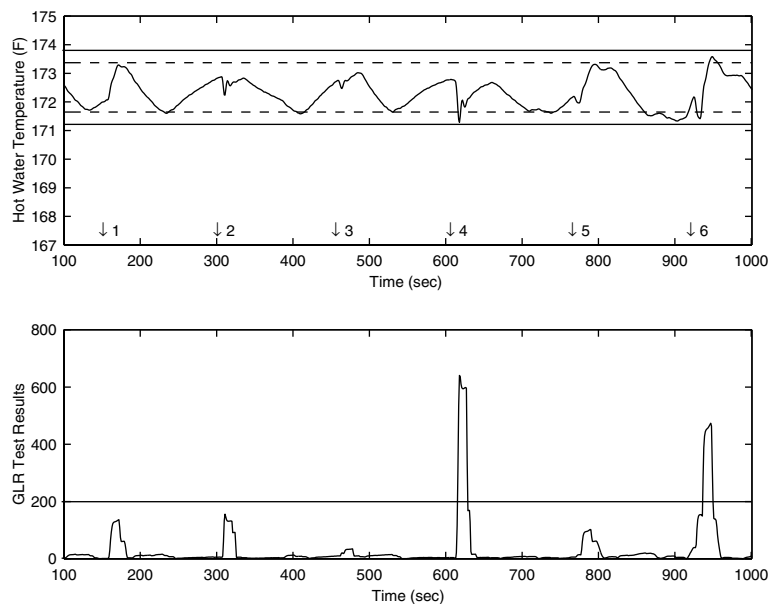

Fig. 4. Holding tube-inlet temperature sensor fault: Hot water temperature measurements with 99% (—) and 95% (- - -) Shewhart confidence limits and GLR test results.

violation and SPM. In case of positive bias in the sensor, the steam valve reduced the steam flow into the system because of the temperature increase. This can cause the hot product temperature to decrease below the critical control limit. Therefore, for proper pasteurization, the diagnosis of the erroneous increase in the holding tube-inlet temperature signal is important. Erroneous reductions in holding tube-inlet temperature readings caused an increase in the steam flow into the system. Therefore, a negative bias in this particular sensor is not a serious threat to product safety.

The time of occurrence of the sensor fault is first depicted in the residuals of that sensor as expected (Figs. 4–9). Since the effect of sensor fault in the controlled variables is propagated in the system through controllers, GLR tests of the actuators eventually showed that there was an abnormality in the system as well. Therefore, if the out-of-control signal is given first in the residuals of controlled variables, it can be deduced that the fault is caused by the corresponding sensors. When the holding tube-inlet temperature and preheater temperature sensors become faulty, inflation in their parity
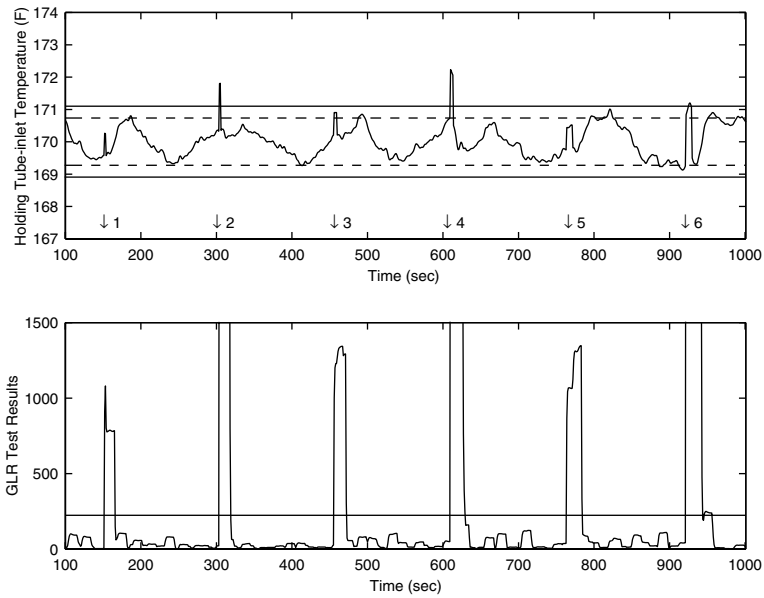
Fig. 5. Holding tube-inlet temperature sensor fault: Holding tube-inlet temperature measurements with 99% (—) and 95% (- - -) Shewhart confidence limits and GLR test results.
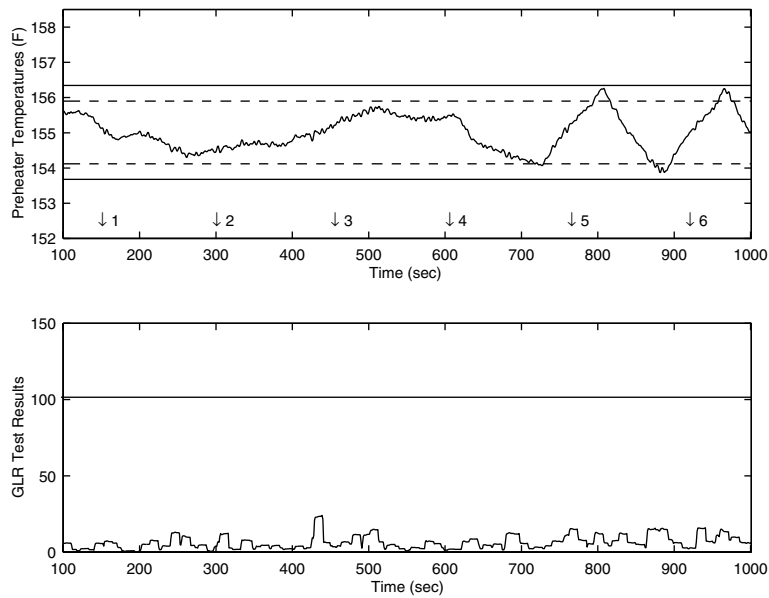

Fig. 6. Holding tube-inlet temperature sensor fault: Preheater temperature measurements with 99% (—) and 95% (- - -) Shewhart confidence limits and GLR test results.

relations is seen first. Then, the false information is carried to controllers, and through the controller responses to the actuators of the system. The actuators change their status, which may lead to deviations in the remaining process variables and eventually to the generation of actuator alarms.

The monitoring and diagnosis charts showed that any deviation in the holding tube-outlet temperature measurements can be detected and used to prevent further process variation by monitoring and regulating the other process variables as well. Since the fluctuations in the product temperature at the exit of the holding tube are caused by the variations or disturbances in the holding tube-inlet temperature sensor, preheater temperature sensor and in the steam valve of the plant, the corrective action can be taken before the fault caused adverse affects on the pasteurized product temperature that will violate the CCP safety limits.
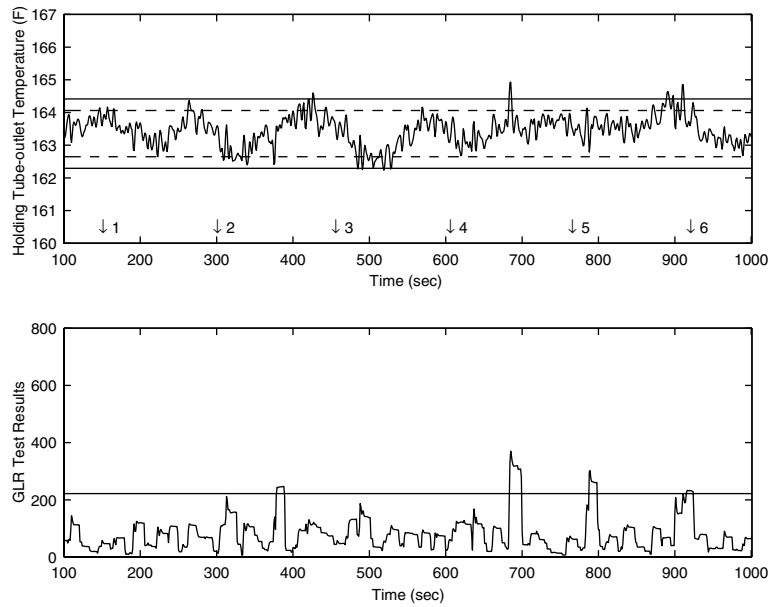
Fig. 7. Holding tube-inlet temperature sensor fault: Holding tube-outlet temperature measurements with 99% (—) and 95% (- - -) Shewhart confidence limits and GLR test results.
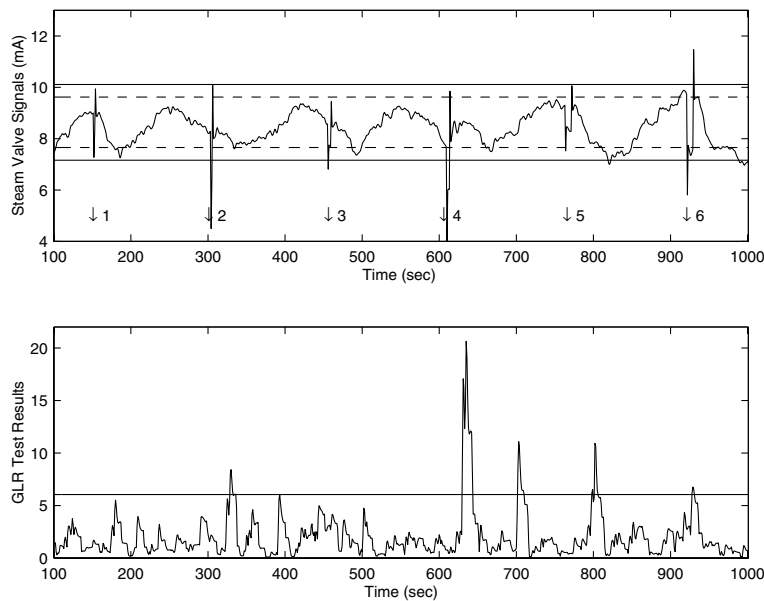

Fig. 8. Holding tube-inlet temperature sensor fault: Steam valve signals with 99% (—) and 95% (- - -) Shewhart confidence limits and GLR test results.

6. Conclusions

Multivariate monitoring and diagnosis techniques have the power to detect unusual events while their impact is too small to cause a significant deviation in any single process variable. This is an important advantage because this trend towards abnormal operation may be the start of a serious failure in the system that may cause failure to properly pasteurize the prod-

uct or lead to equipment damage. The multivariate SPM and diagnosis in food processing operations can play a dual role of quality control and safety assurance by integrating CCP concepts with SPM and FD. Rapid detection of drift towards CCP safety limits by SPM and efficient FD to identify the root cause behind the malfunction in the system saves valuable time. This enables plant operators to take corrective actions quickly when needed before the variation affects significantly the
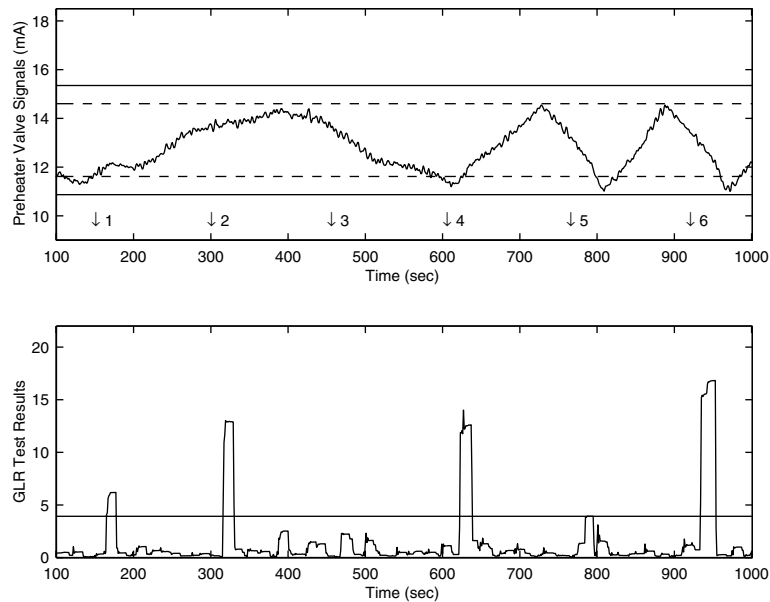
Fig. 9. Holding tube-inlet temperature sensor fault: Preheater valve signals with 99% (—) and 95% (- - -) Shewhart confidence limits and GLR test results.

CCPs of the plant (causing diversion) and trigger drastic interventions that result in production cost increases by stopping process operation or requiring reprocessing of the product.

References

Alanso, C., Acosta, G., Prada, C. D., & Mira, J. (1994). A knowledge-based approach to fault detection and diagnosis in industrial processes: A case study. *Industrial Electronics, IEEE International Symposium*, 397–402.

Bakka, D. (1998). HACCP Programs. *Beverage World, 117*, 98–99.

Bernard, D. (1998). Developing and implementing HACCP in the USA. *Food Control, 9*(2), 91–95.

Bidder, P. L. (1990). Experiences of Introducing SPC in a confectionery factory. *IEE Colloquium on Applied Statistical Process Control*, 1–2.

Buco, S. (1990). How good are your results? An approach to qualitative and quantitative statistical analysis for food monitoring and process control systems. *Food Control, 1*(1), 40–46.

DeCicco, J., Martino, C., Cinar, A., Verdoorn, R., & Balasubramaniam, V. M. (1999). Multivariate statistical monitoring of cooked sausage processes. In *AIChE's Annual Meeting*.

Gertler, J. J. (1998). *Fault detection and diagnosis in engineering systems* (1st ed.). New York: Marcel Dekker, Inc.

Hayes, G. D., Scallon, A. J., & Wong, H. F. (1997). Applying statistical process control to monitor and evaluate the hazard analysis critical control point hygiene data. *Food Control, 8*(4), 173–176.

Hubbard, M. R. (1990). *Statistical Quality Control for the Food Industry* (1st ed.). New York: AVI.

Johnson, R. A., & Wichern, D. W. (1998). *Applied multivariate statistical analysis* (4th ed.). New Jersey: Prentice Hall.

Khandke, S. S., & Mayes, T. (1998). HACCP implementation: A practical guide to the implementation of the HACCP plan. *Food Control, 9*(2), 103–109.

Kosebalaban, F. (2000). Statistical process monitoring and fault diagnosis in a continuous HTST pasteurization process. PhD thesis, Illinois Institute of Technology.

McAnelly, J. K. (1994). HACCP: A total quality system for assuring food safety and quality. *Reliability and Maintainability Symposium, Proceedings, Annual*, 31–36.

Miller, T., & Balch, B. (1991). Statistical process control in food processing. *ISA Transactions, 30*(1), 35–37.

Montgomery, D. C. (1991). *Introduction to statistical quality control* (2nd ed.). John Wiley and Sons.

Motarjemi, Y., Kaferstein, F., Moy, G., Miyogawa, S., & Miyagishima, K. (1996). Importance of HACCP for public health and development, the role of world health organization. *Food Control, 7*(2), 77–85.

Negiz, A., & Cinar, A. (1997). Statistical monitoring of multivariable continuous processes with state space models. *AIChE Journal, 43*, 2002–2020.

Negiz, A., & Cinar, A. (1998). Monitoring of multivariable dynamic processes and sensor auditing. *Journal of Process Control, 8*(5), 375–380.

Negiz, A., Ramanauskas, P., Cinar, A., Schlesser, J. E., & Armstrong, D. J. (1998a). Modeling, monitoring and control strategies for high temperature short time pasteurization systems—1. Empirical model development. *Food Control, 9*(1), 1–16.

Negiz, A., Ramanauskas, P., Cinar, A., Schlesser, J. E., & Armstrong, D. J. (1998b). Modeling, monitoring and control strategies for high temperature short time pasteurization systems—2. Lethality-based control. *Food Control, 9*(1), 17–28.

Negiz, A., Ramanauskas, P., Cinar, A., Schlesser, J. E., & Armstrong, D. J. (1998c). Modeling, monitoring and control strategies for high temperature short time pasteurization systems—3. Statistical monitoring of product lethality and process sensor reliability. *Food Control, 9*(1), 29–48.

Norvilas, A., & Cinar, A. (1997). Multivariate statistical process monitoring in a food processing. Internal Report, Illinois Institute of Technology, Process Monitoring and Control Research Group.

Norvilas, A., Negiz, A., DeCicco, J., & Cinar, A. (2000). Intelligent monitoring by interfacing knowledge-based systems and multivariate statistical monitoring. *Journal of Process Control, 10*, 341–350.

Ozdemir, M., & Ozilgen, M. (1997). Comparison of the quality of hazelnuts unshelled with different sizing and cracking systems. *Journal of Agricultural Engineering Research, 67*, 219–227.

Ozilgen, M. (1998a). Construction of quality control charts with sub-optimal size samples. *Food Control, 9*(1), 57–60.

Ozilgen, M. (1998b). *Food process modeling and control* (1st ed.). Amsterdam: Gordon and Breach Science Publishers.

Peng, Y., Youssouf, A., Arte, P., & Kinnaert, M. (1997). A complete procedure for residual generation and evaluation with application to heat exchanger. *IEEE Transaction on Control System Technology, 5*, 542–555.

Pokkunuri, B. (1994). Knowledge-based simulation for process monitoring and regulatory control. *Intelligent Systems Engineering, 31*, 9–19.

Sanigar, K. M. A. (1990). Statistical process control—the British sugar experience. *IEE Colloquium on Applied Statistical Process Control*, 1–5.

Savage, R. (1995). Hazard analysis and critical control point: A review. *Food Reviews International, 11*(4), 575–595.

Troupis, D., Manesis, S., Koussoulas, N. T., & Chronopoulos, T. (1995). Computer integrated monitoring, fault identification and control for a bottling line. *Industry Applications Conference, IEEE, Thirtieth IAS Annual Meeting, 2*, 1549–1556.

Willsky, A. S., & Jones, H. L. (1976). A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Transactions on Automatic Control* (February), 108–112.