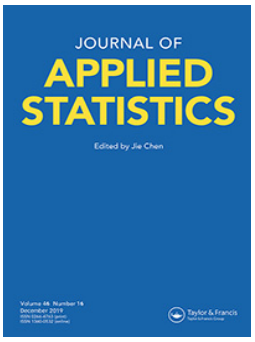

Journal of Applied Statistics

ISSN: 0266-4763 (Print) 1360-0532 (Online) Journal homepage: https://www.tandfonline.com/loi/cjas20

Detecting special-cause variation 'events' from process data signatures

Timothy M. Young, Olga Khaliukova, Nicolas André, Alexander Petutschnigg, Timothy G. Rials & Chung-Hao Chen

To cite this article: Timothy M. Young, Olga Khaliukova, Nicolas André, Alexander Petutschnigg, Timothy G. Rials & Chung-Hao Chen (2019) Detecting special-cause variation 'events' from process data signatures, Journal of Applied Statistics, 46:16, 3032-3043, DOI: 10.1080/02664763.2019.1622658

To link to this article: https://doi.org/10.1080/02664763.2019.1622658

JOURNAL OF APPLIED STATISTICS
2019, VOL. 46, NO. 16, 3032–3043
https://doi.org/10.1080/02664763.2019.1622658

Taylor & Francis
Taylor & Francis Group

APPLICATION NOTE

OPEN ACCESS Check for updates

Detecting special-cause variation 'events' from process data signatures

Timothy M. Young[a], Olga Khaliukova[b], Nicolas André[a], Alexander Petutschnigg[c], Timothy G. Rials[d] and Chung-Hao Chen[e]

[a]Center for Renewable Carbon, University of Tennessee, Knoxville, TN, USA; [b]Assured Bio Labs, LLC, Oak Ridge, TN, USA; [c]Holztechnologie und Holzbau, Salzburg University of Applied Sciences, Kuchl, Austria; [d]Center for Renewable Carbon, University of Tennessee, AgResearch, Knoxville, TN, USA; [e]Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, VA, USA

ABSTRACT

The ability to detect the special-cause variation of incoming feedstocks from advanced sensor technology is invaluable to manufacturers. Many on-line sensors produce data signatures that require further off-line statistical processing for interpretation by operational personnel. However, early detection of changes in variation in incoming feedstocks may be imperative to promote *early-stage* preventive measures. A method is proposed in this applied study for developing *control bands* to quantify the variation of *data signatures* in the context of statistical process control (SPC). *Control bands* based on pointwise prediction intervals constructed from the *Bonferroni Inequality* and Bayesian smoothing splines are developed. Applications using the control band method for data signatures from near-infrared (NIR) spectroscopy scans of industrial fibers of Switchgrass *(Panicum virgatum*) used for biofuels production, Loblolly Pine (*Pinus taeda*) fibers for medium density fiberboard production, and formaldehyde (HCHO) emissions from particleboard were used. Simulations curves (k) of $k = 100$, $k = 1000$, and $k = 10{,}000$ indicate that the Bonferroni method for detecting special-cause variation is closely aligned with the Shewhart definition of control limits when the *pdfs* are Gaussian or lognormal.

ARTICLE HISTORY
Received 26 November 2018
Accepted 16 May 2019

KEYWORDS
Control bands; Shewhart limits; special-cause variation; data signatures; near-infrared spectroscopy

Introduction

Univariate Shewhart control charts have been used extensively by manufacturing and service industries for over 70 years. A plethora of literature is available on SPC, but as noted by Stoumbos *et al.* [1], and quoted by Bischak and Trietsch [2], 'the diffusion of new analytical techniques to application is sometimes slow.' The aim of this paper is to improve the diffusion of research for the practitioner by developing prediction intervals or *control bands* for *data signatures*. The contribution of this work is to illustrate prediction intervals in a new context for real-time interpretation during manufacturing [3]. Rhyne and Treinish [4] appropriately define a 'data signature as a mathematical data vector designed to

CONTACT Timothy M. Young tmyoung1@utk.edu Center for Renewable Carbon, University of Tennessee, 2506 Jacob Drive, Knoxville, TN, USA

characterize a portion of the data set, such as an individual time-frame of a scientific simulation or an article within a corpus.' This paper is aligned well with the spirit of Woodall [5], i.e. ' ... there needs to be a quicker transition from the classical (SPC) methods to newer approaches when appropriate.' A review of the public domain literature indicated a gap in the development of control bands for data signatures for use by practitioners. Given this gap in new SPC methods for data signatures, this study develops control bands for data signatures using NIR spectroscopy data as an example. The well-defined Bonferroni and Bayesian methods for univariate-data are used as a basis for enhanced control band methods for data signatures.

Data challenge

As on-line sensor technologies improve for manufacturing applications, e.g. on-line NIR spectroscopy, X-ray devices, ultrasonic sensors, etc., (see 'Industry 4.0', [6]), operational personnel have an improved ability to assess more advanced real-time process data of feedstocks and intermediate/transformed material characteristics. Many on-line sensors produce full or partial spectral wavelength data (e.g. NIR spectroscopy). It is not always possible to transform this real-time spectral wavelength data into a real-time univariate metric (e.g. PLS model) in a timely manner for operational personnel. This results in the information-loss of real-time data to operational personnel. In this context, developing statistical prediction intervals for data signatures in the *analytical statistical framework* defined by Deming [7, 8] is essential for: (1) quantifying the real-time variation of the process; (2) detecting *special-cause* variation events; (3) maintaining process stability; and (4) timely initiating proactive measures for short-term process improvement. Incoming feedstock variation is problematic for many industries, and reducing such variation is fundamental to lowering operational targets, e.g. weight, density, bonding agents, drying time, energy usage, etc. [9]. The basic theory of entropy (second law of thermodynamics) implies that processes tend toward a state of disorder. If something appears to be decreasing in entropy, the component may be influencing another component in the system, and causing this next component to increase its entropy, e.g. weight variation, bulk density variation, process cycle time changes, etc. Statistical process control (SPC) is an excellent tool for detecting real-time entropy. Focusing on variation reduction which ultimately leads to cost reduction is essential for ensuring business success [3, 8, 10, 11].

Statistical methods have played key roles in the improvement process philosophies for manufacturing industries [5], e.g. *Six Sigma Quality, Lean Six Sigma (LSS), continuous improvement, Total Quality Management (TQM), continuous improvement, etc.* Many *enumerative* and *analytical* statistical methods exist for quality improvement through the quantification and understanding of sources of variation [5, 11–13]. Deming [7] urged us to distinguish between *enumerative* and *analytical* studies. *Enumerative studies* deal with characterizing an existing, finite, unchanging target population by sampling from a well-defined frame, e.g. ANOVA, Design of Experiments, confidence intervals, etc. [12]. In contrast, *analytical studies* most frequently encountered in industrial applications, focus on real-time analysis of a process or system with the aim of process improvement and prediction, e.g. prediction intervals, statistical process control (SPC), control charts, etc. [7, 14]. A key tool in the application of SPC is the implementation of Shewhart and other types of control charting techniques to quantify and detect variation in the process. Predicting outcomes of manufacturing applications from such charts is fundamental to avoiding scrap and rework of the final product [1, 15, 16].

As Ceriolo *et al.* [17] noted: 'Functional data analysis (FDA) concerns the statistical analysis of data which come in the form of continuous functions, usually smooth curves.' In FDA, each data signature is seen as a single entity, rather than a collection of individual observations [5, 17]. As Morris *et al.* [18] noted, 'methods that model functional profiles in their entirety have the potential to extract more information from the data compared with methods based on arbitrary chosen summary measures,' i.e. univariate methods may be inappropriate for detecting real-time changes in quality characteristics that are inherently non-univariate, [3, 19–25].

Relevant data

NIR spectroscopy data from scans of Switchgrass fiber (*Panicum virgatum*) as a feedstock ($k = 41$) for the production of biofuels are provided by Genera Energy (www.spc4lean.com), where k = *number of data signatures or curves*. The NIR data of real-time scans of Switchgrass feedstocks without post-statistical analyses are a good example of the application of control bands to assess the natural- and special-cause variation of data signatures. The second data are from NIR spectroscopy scans of core fiber samples of Loblolly Pine (*Pinus taeda*) used for medium density fiber (MDF) production ($k = 100$). The second data were derived from a related study to measure characteristics of the Loblolly pine fibers as related to mechanical properties of the final product [26]. The third data are from NIR spectroscopy scans of formaldehyde (HCHO) emissions from particleboard after the pressing stage ($k = 20$).

All three data sets were used to quantify the natural variation of the data signatures and detect the possible special-cause variation. Robustness of the proposed control band methods was tested using simulation of data signatures assuming the common Gaussian, lognormal, and Weibull *pdfs* [9, 27]. The simulation data signatures were performed for three sets of data signatures, i.e. $k = 100$, $k = 1000$, and $k = 10{,}000$. The proposed approach is important for providing a methodology for developing statistical prediction intervals for data signatures to improve real-time information for manufacturing operational personnel.

Methods

Bonferroni methods as control bands

One approach to developing prediction intervals for data signatures is based on enhancing the fundamentals presented by Bonferroni inequality for univariate data [28]. A general form of the Bonferroni inequality as noted by Milton and Arnold [29] is *let* $A_1, A_2, \ldots, A_c$ *be events then,*

$$P[A_1 \cap A_2 \cap \ldots\ldots \cap A_c] \geq 1 - [P[A_1{}'] + P[A_2{}'] + \ldots\ldots + P[A_c{}']] \qquad (1)$$

As applied to the study of statistical intervals, the Bonferroni bound from elementary probability theory provides a simple, conservative lower bound on the actual α-level for a joint interval-statement [14]. If the simultaneous intervals are statistically independent, the joint α-level is,

$$(1-\alpha_J) \geq 1-\alpha_1-\ldots\ldots\alpha_K \tag{2}$$

As Hahn and Meeker [14] note, Equation (2) 'provides a useful way of combining interval statements to give a conservative bound for the actual joint α-level.' Fisher [30] called this a 'confidence ribbon' since the pointwise statistical intervals are extended until they have the desired simultaneous coverage probability of 1 α. Hahn and Meeker [14] provided a two-sided 100(1 α)% simultaneous prediction interval to contain the values of all of m future randomly selected observations from a previously sampled population (or process) that can be described by a normal distribution,

$$[y_{IB}, y_{UB}] = \bar{x} \pm r_{(1-\alpha;\mu,v)}s \tag{3}$$

where $[y_{IB}, y_{UB}]$ is a two-sided statistical interval, $\bar{x}$ is the process average, s is the process standard deviation and $r_{(1-\alpha;\mu,v)}$ is the factor for calculating a normal distribution two-sided 100(1 α)% prediction interval for m future observations using the results of a previous sample of n observations. A conservative approximation for $r_{(1-\alpha;\mu,v)}$ is,

$$r_{(1-\alpha;m,n)} \cong \left(1+\frac{1}{n}\right)^{\frac{1}{2}} t_{((1-\alpha)/(2m);n-1)^2} \tag{4}$$

where $t_{((1-\alpha)/(2m);n-1)} \approx z_{(1-\alpha)/(2n)}$ [14]. Thus, normal distribution percentiles provide a generally adequate approximation for t distribution percentiles when n is large and $1-\alpha/2$ is not too large (e.g. $t_{(0.975,60)} = 2.000$ and $z_{(0.975)} = 1.960$).

In the spirit of the Bonferroni method, Hardle [28] proposed constructing pointwise prediction intervals on k observations at each value of x. The form of the simultaneous prediction intervals for any data signature as proposed by Hardle [28] using the Bonferroni method is,

$$\bar{y}_i \pm z_{(1-\alpha)/(2n)} s_i \left(1+\frac{1}{k}\right)^{\frac{1}{2}} \tag{5}$$

for $i = 1, \ldots, n$, $\bar{y}_i$ is the average curve of k observed curves, $z_{(1-\alpha)/(2n)}$ is the $1-\alpha$ percentile for the standard normal *pdf*, s_i is standard deviation across the k curves Equation (5) is the approach followed in this study.

Bayesian splines as control bands

In many applications, smooth shapes without discontinuities are preferred. Smoothing spline curves are useful in applications that require nonparametric regression models. Wahba [31] proposed Bayesian confidence intervals for a smoothing spline. As noted by Wang and Wahba [32] it is highly desirable to have interpretable statistical intervals for these nonparametric estimates. Wang and Wahba [32] noted 'that the best variations of the bootstrap intervals behave similar to the Bayesian intervals.' Consider the model proposed by Wahba [31]:

$$y_i = f(t_i) + \varepsilon_i \tag{6}$$

where, $i = 1, \ldots, n$, $t_i \in [0, 1]$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^T \sim N(0, \sigma^2)$, σ^2 unknown, and $f \in W_m$ where,

$$W_m = \left\{ f : f, f, \ldots, f^{(m-1)} \text{ absolutely continuous}, \int_0^1 \left(f^{(m)}\right)^2 dt < \infty \right\} \tag{7}$$

Generally $f(t)$ is a piecewise polynomial, e.g. a cubic spline has the following form over $[i, i+1]$,

$$y(t) = a_y t^3 + b_y t^2 + c_y t + \varepsilon \tag{8}$$

The smoothing spline f_λ minimizes

$$\frac{1}{n} \sum_{i=1}^{n} (\psi_l - \phi(\tau_l))^2 + \lambda \int_0^1 (f^{(m)})(t))^2 dt \tag{9}$$

over $f \in W_m$. The amount of local averaging, and therefore the smoothness of the estimator, is controlled by the value of λ. Given Equations (7), (8), and (9), Wahba [31] proposed the $(1-\alpha)100\%$ Bayesian confidence intervals for $\{f(t_i)\}_{i=1,n}$ as,

$$f_\lambda(t_i) \pm z_{\alpha/2} s \sqrt{h_i} \tag{10}$$

where $i = 1,.., n$, $z_{\alpha/2}$ is the $1-\alpha/2$ percentile of a standard normal distribution, s is an estimate of σ, and h_i is the i^{th} diagonal element of the *hat* matrix. Equation (10) is the second approach examined in this study.

Application

Applications of the control bands using the Bonferroni method detected special-cause variation for the data signatures ($k = 41$) from Switchgrass *(P. virgatum)* as illustrated in Figures 1 and 2.

The NIR spectral wavelengths detected one data signature that is completely out-of-control below the lower control limit near the wavelengths 1290–1380 nm, and another that had portions of the data signature out-of-control near wavelengths of 1300–1330 nm. Detecting such events in a real-time setting will alert operational personnel that the feedstock quality is changing and be a proactive first step for more detailed investigation of the process, e.g. change in the age of feedstock, moisture of feedstock, contamination, etc. The control bands also quantify the natural- or common-cause variation of the data signatures from incoming feedstocks and may also avoid over-adjustment of the process from single curve assessment, e.g. see Deming's [8] *Funnel Experiments* related to over-adjusting processes by operational personnel.

Special-cause data signature variation for Loblolly pine (*P. taeda*) feedstocks is also detected from the NIR spectroscopy scans using the Bonferroni method for prediction

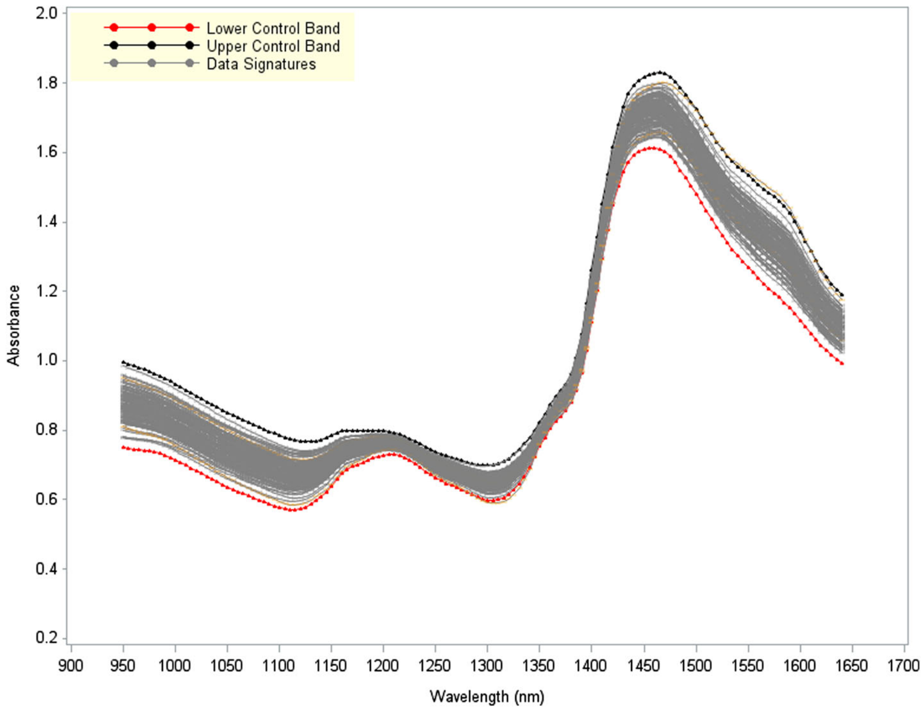

Figure 1. Bonferroni control bands for spectral data signatures of Switchgrass (*P. virgatum*) fiber feedstocks for biofuels.

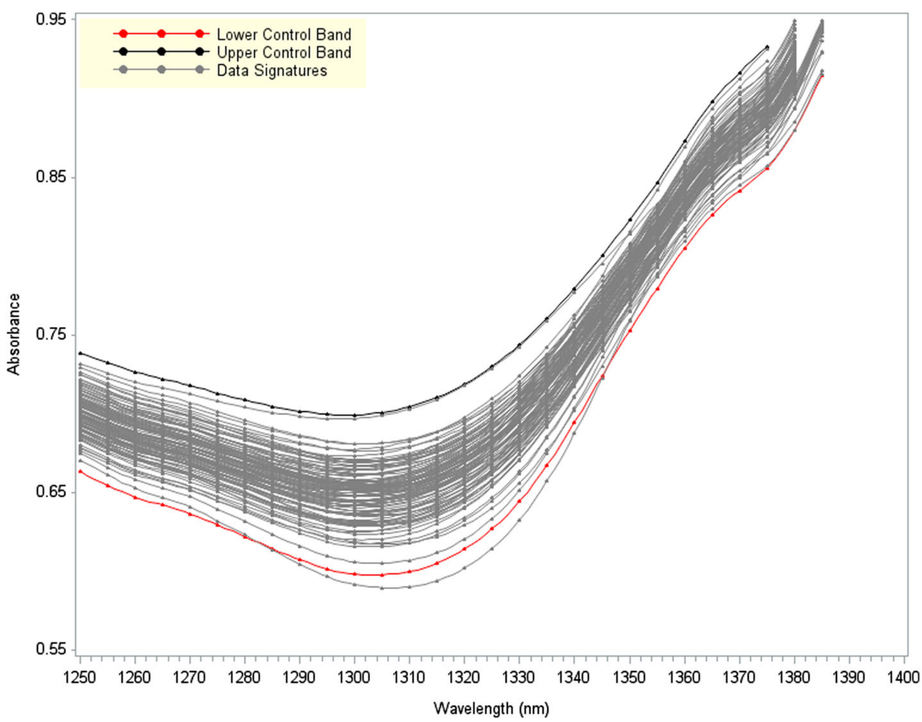

Figure 2. Zoomed-in section from Figure 1 for Switchgrass (*P. virgatum*) fiber feedstocks for biofuels.

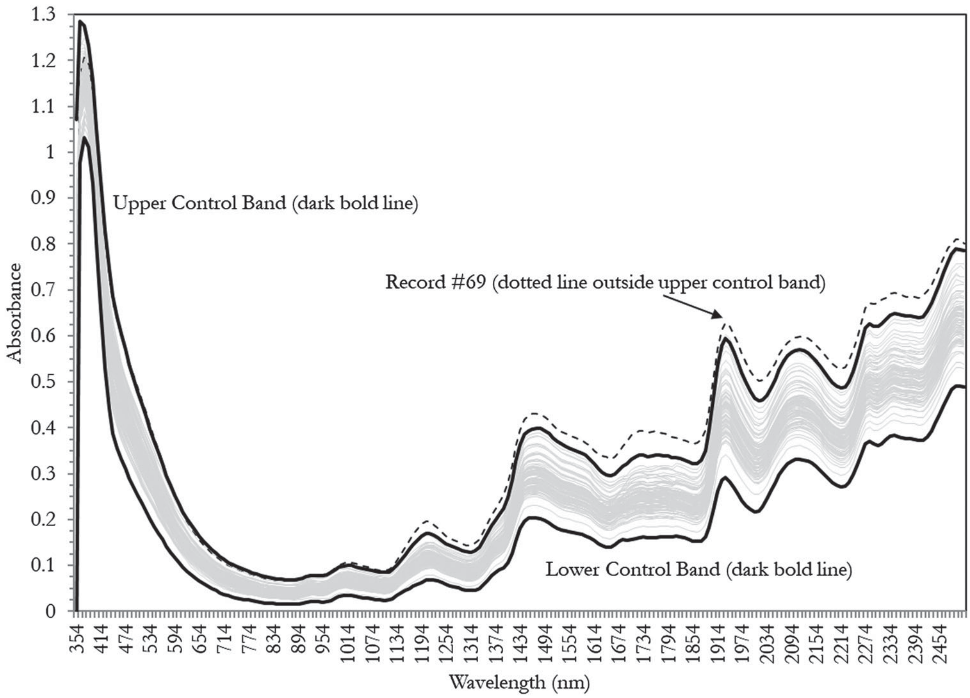

Figure 3. Bonferroni control bands for spectral data signatures of Loblolly pine (*P. taeda*) fiber feedstocks for medium density fiberboard (MDF).

intervals (Figure 3). See the strong signal of out-of-control for the data signature #69 above the upper control limit highlighted in Figure 3. Such a strong signal is an alert for immediate investigation.

Out-of-control signals of formaldehyde (HCHO) emissions are detected in data signatures #17 and #20 from the NIR spectroscopy scans of particleboard after pressing (Figures 4, 5 and 6). Even though the signatures vary by wavelength (*nm*) location, HCHO emissions are regulated by the US government not to exceed a limit of 0.10 ppm, which if exceeded requires a quarantine of final product. Current test methods for estimating HCHO emissions are conducted by periodic off-line destructive lab tests (e.g. taken one to two hours apart). Such tests take several hours before results are relayed to operational personnel. Detection of real-time special-cause variation for HCHO emissions could trigger an alert for closer monitoring of the process and prompt additional lab tests to ensure legal product conformance.

The figures of the control bands illustrate the importance of this methodology for practitioners for detecting real-time events that are absent from current manufacturing analytics, i.e. operational personnel typically does not view data signatures from sensors in the context of statistical prediction intervals. If presented with such intervals, operational personnel are immediately alerted to a change in the quality characteristics of the incoming feedstocks and can take appropriate precautionary measures. This prevents the manufacture of reject or off-grade product which is closely aligned with the Shewhart philosophy of detecting instability in processes [15].

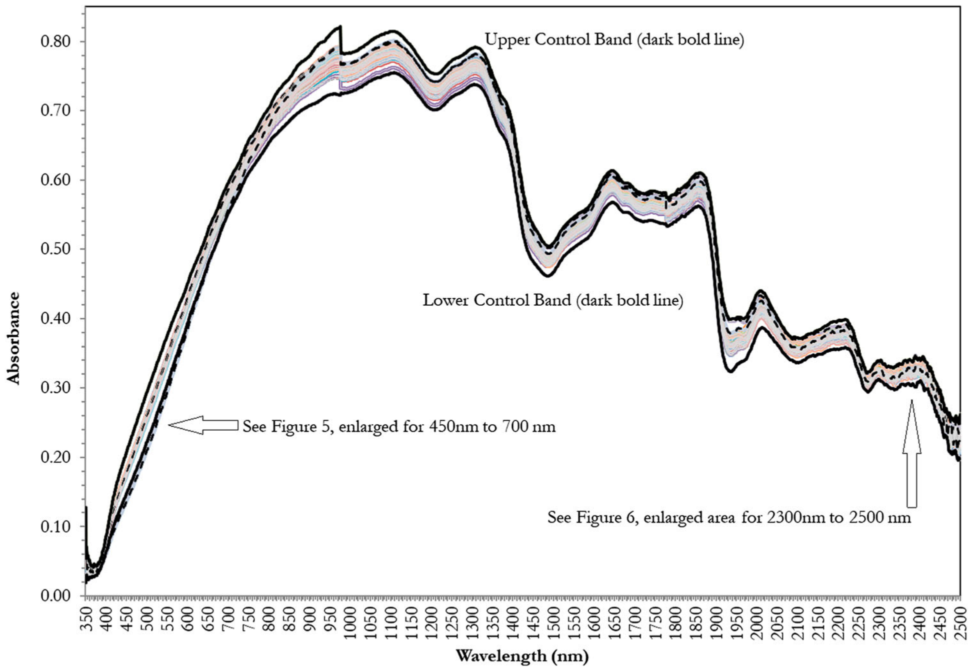

Figure 4. Bonferroni control bands for spectral data signatures of HCHO emissions of final particleboard production.

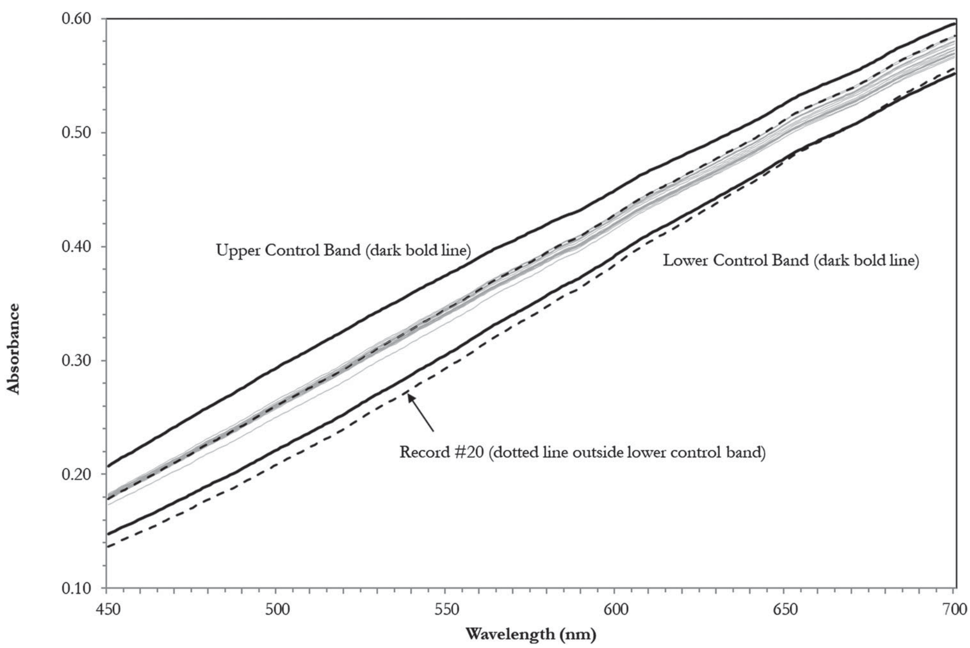

Figure 5. Zoomed-in section for Figure 4 for spectral data signatures of HCHO emissions (450–700 *nm*) of final particleboard production.

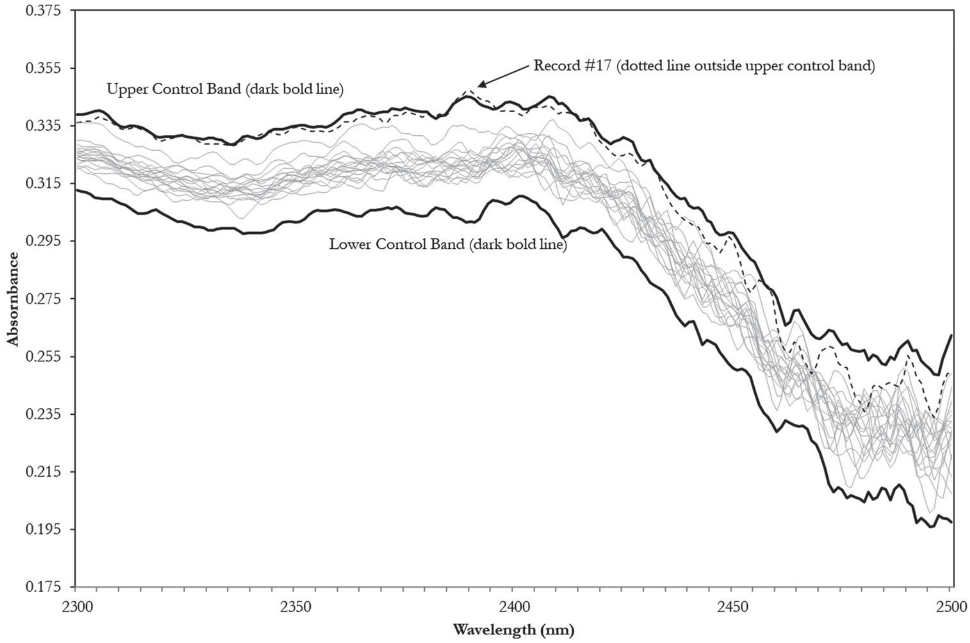

Figure 6. Zoomed-in section for Figure 4 for spectral data signature #17 of HCHO emissions (2300–2500 *nm*) of final particleboard production.

Simulated data signatures

To test the 'false alarm' rates for the control bands, simulations were developed for data signatures assuming the Gaussian, lognormal and Weibull *pdfs* which are common to many industrial applications [27, 33]. Shewhart [15] assumed a Gaussian or normal *pdf* for univariate data in the development of Shewhart control limits. Given that Shewhart [15] developed analytical statistical methods for real-time prediction intervals that are approximately $\bar{X} \pm 3s$, the expected false alarm rate for Shewhart's univariate data applications is 0.003. This same false alarm rate is applied to the control band methods of this study. The results based on the simulated curves for each *pdf* and k curves are given in Table 1.

The matched results for the Bonferroni and Spline methods for a Gaussian *pdf* were consistent with the Shewhart false alarm rate. The Bonferroni method was slightly lower in false alarm rate by two curves for $k < 1000$; and the spline method was higher by two or one curves for $k < 1000$. The difference for both methods relative to the Shewhart method in alarm rate was ± 2 curves for $k > 10{,}000$. The small differences in false alarm rates are a useful result in the detection of special-cause variation and are consistent with Shewhart limits. This was also true as the *pdfs* departed from the Gaussian as tested under the assumptions of lognormal or Weibull *pdfs*. The Bonferroni method when the *pdf* was Weibull was closer aligned to Shewhart false alarm rates. This control band method is important in applications since many random variables as related to the strength of materials are inherently Weibull. The Bayesian spline smoothing method had slightly higher false alarm rates for the Gaussian and lognormal *pdfs* (Table 1). However, the false alarm rate was identical to the Shewhart rate for the Weibull *pdf*. The ability to quantify and distinguish

Table 1. Count of randomly simulated curves for the Gaussian, lognormal and Weibull parameters that are beyond the estimated control bands.

	Shewhart control chart			Bonferroni method			Bayesian spline smoothing method		
Curves	*Gaussian*	*Lognormal*	*Weibull*	*Gaussian*	*Lognormal*	*Weibull*	*Gaussian*	*Lognormal*	*Weibull*
100	4	2	7	2	2	4	5	4	7
1000	7	4	11	5	3	7	9	8	10
10000	47	37	52	46	34	48	49	40	52

between natural (common-cause) and special-cause variation in a real-time operational setting may prevent the manufacture of the defective product and reduce future warranty claims.

Conclusions

The ability to quantify the real-time variation of data signatures in manufacturing is important. By quantifying the natural (common-cause) variation of the process and distinguishing such variation from special-cause variation (events), operational personnel have the ability to monitor real-time process stability from advanced data. Given that applications of advanced sensor technologies for the rapid assessment of quality on production lines are increasing and produce non-univariate data signatures, the results of this study may be a useful step in advancing the applications of SPC. Of the two control band methods examined, the Bonferroni control band method was closely aligned with the Shewhart philosophy when the *pdf* is either Gaussian or lognormal. The Bayesian spline smoothing method may be better when the *pdf* is Weibull.

Advancing SPC methods that quantify process variation from on-line sensor technology is essential for the success of *Industry 4.0* and the advancement of real-time analytics. As operational personnel in manufacturing progress as data scientists, development of new SPC methods for data signatures is essential for sustaining a competitive business advantage. A precursor to reducing variation in a process is to first quantify variation. Variation reduction leads to target size reduction, lower costs, and avoidance of warranty claims.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Cooperative State Research, Education, and Extension Service [grant number TEN00MS-107, Young]; U.S. Department of Energy [grant number R11-3215-096].

References

[1] Z.G. Stoumbos, M.R. Reynolds Jr, T.P. Ryan, and W.H. Woodall, *The state of statistical process control as we proceed into the 21st Century*. J. Am. Stat. Assoc. 95 (2000), pp. 992–998.
[2] D.P. Bischak, D. Trietsch, *The rate of false signals in control charts with estimated limits*. J. Qual. Technol. 39 (2007), pp. 54–65.
[3] W. Woodall, H. Spitzner, D.J. Montgomery, et al., *Using control charts to monitor process and product quality profiles*. J. Qual. Technol. 36 (2004), pp. 309–320.

[4] T.M. Rhyne, L. Treinish, *Data signatures and visualization of scientific data sets*. IEEE Comput Graph 3/4 (2000), pp. 12–15.
[5] W.H. Woodall, *Controversies and contradictions in statistical process control*. J. Qual. Technol. 32 (2000), pp. 341–350.
[6] Y. Liao, F. Deschamps, E. de Freitas, et al., *Past, present and future of Industry 4.0 - a systematic literature review and research agenda proposal*. Int. J. Prod. Res. 55 (2017), pp. 3609–3629.
[7] W.E. Deming, *On probability as a basis for action*. Am. Stat. 29 (1975), pp. 146–152.
[8] W.E. Deming, *Out of the Crisis*, Massachusetts Institute of Technology, Center for Advanced Engineering Study, Cambridge, MA, 1986.
[9] G. Taguchi, *Taguchi on Robust Technology Development*, The American Society of Mechanical Engineers, American Society of Mechanical Engineers (ASME) Press, New York, NY, 1993.
[10] D.J. Edwards, R.V. León, T.M. Young, F.M. Guess, and K.A. Crookston, *Comparison of two wood plastic composite extruders using bootstrap confidence intervals on measurements of sample failure data*. Qual. Eng. 25 (2013), pp. 23–33.
[11] J.F. Lawless, R.J. Mackay, J.A. Robinson, *Analysis of variation transmission in manufacturing processes – part I*. J. Qual. Technol. 31 (1999), pp. 131–142.
[12] G.J. Hahn, *Deming's impact on industrial statistics: some reflections*. Am. Stat. 49 (1995), pp. 336–341.
[13] M.R. Reynolds Jr, Z.G. Stoumbos, *Monitoring the process mean and variance using individual observations and variable sampling intervals*. J. Qual. Technol. 33 (2000), pp. 181–205.
[14] G.J. Hahn, W.Q. Meeker, *Statistical Intervals*, John Wiley and Sons, New York, NY, 1991.
[15] W. Shewhart, *Economic Control of Quality of Manufactured Product*, D. Van Nostrand Company, New York, NY, 1931.
[16] T.M. Young, N.E. Clapp Jr, F.M. Guess, and C.-H. Chen, *Predicting key reliability response with limited response data*. Qual. Eng. 26 (2014), pp. 223–232.
[17] A. Ceriolo, F. Laurini, and A. Corbellini, *Functional cluster analysis of financial time series*, Proc. of the Classification and Data Analysis Group of the Italian Statistical Society, September 22–24, Bologna, Italy, 2003.
[18] J.S. Morris, C. Arroyo, B.A. Coull, L.M. Ryan, R. Herrick, and S.L. Gortmaker, *Using wavelet-based functional mixed models to characterize population heterogeneity in accelerator profiles: a case study*. J. Am. Stat. Assoc. 101 (2006), pp. 1352–1364.
[19] E. Bertran, M. Blanco, S. Maspoch, M.C. Ortiz, M.S. Sanchez, and L.A. Sarabia, *Handling intrinsic non-linearity in near-infrared reflectance spectroscopy*. Chemomet Intell. Lab. 49 (1999), pp. 215–224.
[20] A. Bowman, S. Young, *Graphical comparison of nonparametric curves*. Appl. Stat. 45 (1996), pp. 83–98.
[21] B.M. Colosimo, Q. Semeraro, M. Pacella, *Statistical process control for geometric specifications: on monitoring of roundness profiles*. J. Qual. Technol. 40 (2008), pp. 1–18.
[22] S. Golowich, J. Landwehr, S. Vander Wirl, *Interplay between physics and statistics for modeling optical fiber bandwidth*. Technometrics. 44 (2002), pp. 215–229.
[23] N. Heckman, R.H. Zamar, *Comparing shapes of regression functions*. Biometrika 87 (2000), pp. 135–144.
[24] K. Kim, M.A. Mahmoud, W.H. Woodall, *On the monitoring of linear profiles*. J. Qual. Technol. 35 (2003), pp. 317–328.
[25] M.S. Sanchez, E. Bertran, L.A. Sarabia, M.C. Ortiz, M. Blanco, and J. Coello, *Quality control decisions with near infrared data*. Chemometr. Intell. Lab. 53 (2000), pp. 69–80.
[26] T.G. Rials, S.S. Kelley, C.-L. So, *Use of advanced spectroscopy techniques for predicting the mechanical properties of wood composites*. Wood Fiber Sci. 34 (2002), pp. 398–407.
[27] C.-D. Lai, D.N. Murthy, M. Xie, *Weibull distributions and their applications*, in *Springer Handbook of Engineering Statistics*, Pham H., ed., Springer Handbooks, Springer, London, 2006.
[28] W. Hardle, *Applied Nonparametric Regression*, Cambridge University Press, New York, NY, 1990.

[29] J.S. Milton, J.C. Arnold, *Introduction to Probability and Statistics – Principles and Applications for Engineering and the Computing Sciences*, Fourth Ed, McGraw Hill, Inc., Boston, MA, 2003.

[30] N. Fisher, *Graphical methods in statistics: current and prospective views*, Proc. of the 46th ISI Session 3, 1987.

[31] G. Wahba, *Bayesian confidence intervals for the cross-validated smoothing spline*. J. R. Stat. Soc. 45 (1983), pp. 133–150.

[32] Y. Wang, G. Wahba, *Bootstrap confidence intervals for smoothing splines and their comparison to Bayesian confidence intervals*. J. Stat. Comput. Sim. 51 (1995), pp. 263–279.

[33] K.E. Vokovinsky, L.B. Pfahler, *The role of normal data distribution in pharmaceutical development and manufacturing pharmaceutical technology*. Pharm. Tech. 38 (2014), pp. 1–3.