Available online at www.sciencedirect.com

ScienceDirect

Manufacturing Letters 1 (2013) 25–27

www.elsevier.com/locate/mfglet

Research Letters

Estimating spatial and temporal patterns of defects

Matthew Plumlee, Jianjun Shi *

Industrial and Systems Engineering, Georgia Institute of Technology, 765 Ferst Drive, Rm 109, Atlanta, GA 30332, USA

Received 1 September 2013; accepted 27 September 2013
Available online 3 October 2013

Abstract

This work investigates the use of stochastic modeling and statistical estimation to study patterns of defects in or on products. We conjecture about underlying defect patterns by estimating the unknown intensity function based on the observed defects. This work does not assume the patterns follow a parametric form and therefore the described estimate is flexible enough to be used in many different production settings. Details for estimation are provided and an example is presented using data collected from a steel rolling process.

Keywords: Inhomogeneous process; Point process; Surface quality; Statistical process control; Steel rolling

1. Problem description

Automated detection of defects on products has become a reality thanks to developments in video recording, computer processing and digital storage technology. After detection, the spatial coordinate and/or occurrence time of each defect is stored for analysis. Examples of production environments where this type of measurement is used include:

- Silicon wafers, the basis for semiconductors and many types of solar cells, should be free of particulate matter for downstream processing. Investigations have been performed via an optical inspection device that can detect particles on each wafer [1–3].
- Nanocomposites are formed by infusing nanoscale particles into existing materials [4]. Clustered particles will abate the advantages of the composite over the original material. The size and location of clusters has recently been studied [5,6].
- Steel rolling, used to form many different steel parts, can introduce surface level defects that harm the performance of final product. Using specially designed cameras, methods have been developed to accurately detect seam formation and other surface defects [7,8].

The presence of some defects is almost always unavoidable due to the production mechanisms used. A more pressing concern is the underlying *cause* of the defects, which is often indicated by their pattern. This paper is designed to answer the following question: How could one compile the data to give a visual indication of where or when defects are more likely? When responding to this question, the variability in the underlying process must be accounted for. Using stochastic modeling, this paper answers this question through a mathematical lens.

Despite the prevalence of this type of data and the generality of the question at hand, the methods used to analyze these systems are not fully developed. The simple and popular technique of binning [1–3,6], where the number of defects in each region is saved and monitored, has recently been shown to be an ineffective strategy [9]. Binning is not a powerful method because the storage of data into binned counts represents a destruction of the continuously indexed information. New tools are needed

---

* Corresponding author. Tel.: +1 404 385 3488.
E-mail addresses: mplumlee@gatech.edu (M. Plumlee), jianjun.shi@isye.gatech.edu (J. Shi).

2213-8463/$ - see front matter
http://dx.doi.org/10.1016/j.mfglet.2013.09.006

to analyze these systems and this paper discusses a promising strategy that does not employ binning.

2. Mathematical background

To formalize notation, say we observe n products with m total defects. Each product is indexed by i and the number of defects on the ith product is denoted m_i and $\sum_{i=1}^{m} m_i = m$. The defect locations for the ith product are labeled as $x_{i1}, \ldots, x_{im_i}$. The location x is defined generally, it can represent a time, a one-dimensional spatial coordinate, a multidimensional coordinate, or a time-space coordinate vector. The locations x are located in some space X, e.g. a given production run period or the surface of a product. Stochastic models that produce a series of points are known as point process models [10].

Let $f(x)$ represent the *intensity function* of the surface. The intensity function is defined as follows: given a subset, $X_0 \subset X$, the expected number of defects in that region is given by $\int_{X_0} f(x)\,\mathrm{d}x$. The intensity function indicates the rate at which faults occur in X. Figure 1 illustrates examples of one-dimensional intensity functions and 10 draws from those processes. A flat intensity function generates evenly spread defects. Peaks in the intensity function correspond to regions with high rates of defects.

Using this framework, the original objective of the paper can be restated with a mathematical equivalent: What is a good estimate of the underlying intensity function based on the observed defects?

3. Proposed methodology

The exact methods for estimating $f(x)$ are tied to the underlying assumptions on $f(x)$. We focus our attention to nonparameteric models of f, which are developed without restrictive assumptions on the form of f. For example, the form of $f(x) = \beta_0 + \beta_1 x$ is a linear parametric model with parameters β_0 and β_1. This form of $f(x)$ can be used to estimate intensity functions that fit this exact formula, but is incapable of modeling general shapes of defect patterns.

Historically, the most prevalent methods of intensity function estimation (e.g. [11–13]) are variations on locally smoothed estimates. This method is nonparametric in two senses. First, it is not limited to a particular type of point process (e.g. Poisson process). Second, it does not assume the function can only take on a limited number of shapes, e.g. linear parametric form.

Before we outline the locally smoothed estimate, we first discuss the simplified case when $f(x) = \alpha$, where α is a constant. An estimate of f based on the observed defects is computed by the maximizing criteria

$$L(f) = \sum_{i=1}^{n}\sum_{j=1}^{m_i} \log\alpha - n\alpha|X|,$$

with respect to α, where $|X| = \int_X \mathrm{d}x$. The optimal value of α implies that

$$\hat{f}(x) = m/n.$$

However, this estimate is not useful in the majority of cases as $f(x)$ is often not a constant. Alternatively, the locally smoothed estimate [14] is found by using a kernel function $K_y(|x-y|)$ with two properties: it should be monotonically decreasing away from 0 and $\int_X K_y(|x-y|)\mathrm{d}x = 1$ for all $y \in X$. The estimate is then given by

$$\hat{f}(x) = \underset{\alpha_x}{argmax} \sum_{i=1}^{n}\sum_{j=1}^{m_i} K_{x_{ij}}(|x-x_{ij}|)\log\alpha_x - n\alpha_x \int_X K_x(|x-z|)\mathrm{d}z,$$

which is a 'local estimate' because more weight being placed on observations near x. After some derivations, one can establish

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m_i} K_{x_{ij}}(|x-x_{ij}|).$$

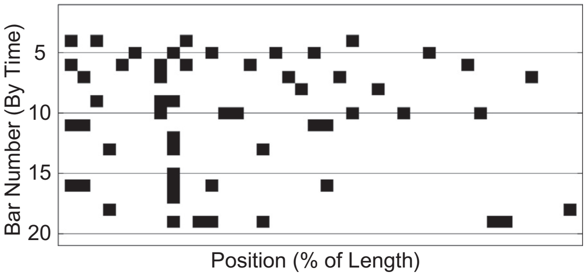

Figure 2. Examples of defects observed during a rolling process, a black box indicates at least one defect in the area.

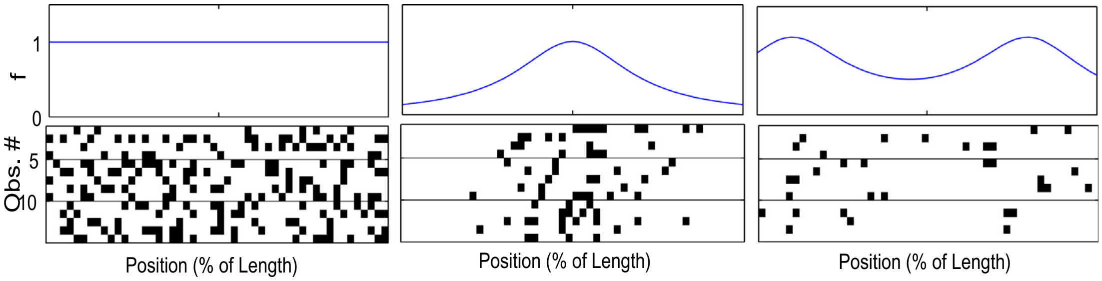

Figure 1. Examples of intensity functions (top) and point processes generated by inhomogenous Poisson process models with corresponding intensity functions (bottom). In the bottom figures, a black box represents at least one defect in that location.

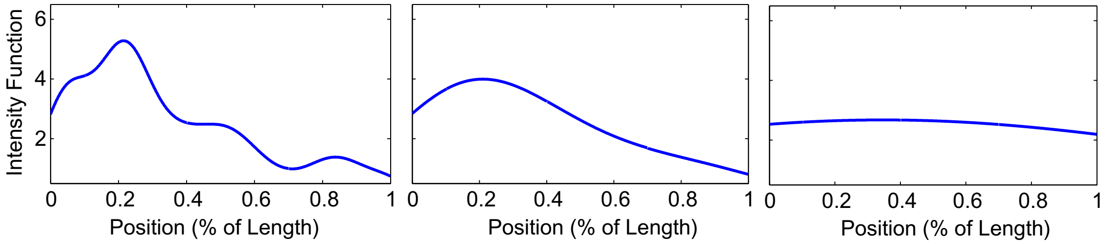

Figure 3. Examples of estimated intensity functions using the data in Figure 2 with $\theta = 15$ (left), 5 (middle), and 1 (right).

The value of this estimate is two-fold: (1) no numerical optimization is needed, so the estimate is fast to compute and (2) it is a localized estimate that can be close to a broad array of functions $f(x)$.

4. Example

This section details an example that will hopefully fix the ideas discussed in this work. The production data in Figure 2 was collected from a steel rolling process, where x is the longitudinal coordinate of the defect in $[0,1] = X$ (the normalized length). A total of $n = 19$ products are observed. Let

$$K_y(|x-y|) = \frac{\theta\sqrt{2\pi}}{\Phi(\theta y) - \Phi(\theta y - \theta)} \exp(-\theta^2|x-y|^2),$$

which is known as the Gaussian kernel where Φ is the cumulative standard normal distribution. It is easy to verify this satisfies both the monotonically decreasing and unit integrality conditions, so it is a valid kernel function.

An important practical concern is how to select θ , which controls the lengthscale in the basis functions used in the estimate. A visual method is done by first computing $\hat{f}(x)$ using several different values of θ and selecting the most appealing estimate, see Figure 3. The value of θ can also be selected by cross validation. To do this, first fit $\hat{f}_{CV}(x)$ using the first 14 observations. The remaining 5 observations are used to compute the cross-validation score,

$$T_{CV} = \sum_{i=14}^{19}\sum_{j=1}^{m_i} \log\{\hat{f}_{CV}(x_{i,j})\} - 5\int_X \hat{f}_{CV}(z)\mathrm{d}z,$$

and we would like to maximize T_{CV} with respect to θ. The values of T_{CV} for $\theta = 15$, 5 and 1 (seen in Figure 3) are -0.7468, -0.3321, and -0.4936 respectively, which implies we should use the estimate with $\theta = 5$.

5. Discussion

This paper describes a method of estimating the underlying pattern of defects via the use of inhomogeneous point process models. The prevalence of defects being stored as point data has grown immensely over the past 20 years and the trend appears to be accelerating. As sensing equipment becomes more sophisticated, industrial production environments will require powerful methods to analyze these types of data. This paper does not attempt to address all of the concerns of a production engineer. More research is needed to investigate challenging problems such as defect pattern monitoring, discovering inaccurately identified defects and processing huge pools of defect information. Additionally, the problem of translating an estimated defect pattern to the underlying root cause is an important concern in implementation that is not addressed in this work.

Acknowledgments

This research was supported by the National Science Foundation grant CMMI-1030125. The data used in Section 4 was provided by OG Technologies, Inc.

References

[1] Taam W, Hamada M. Detecting spatial effects from factorial experiments: an application from integrated-circuit manufacturing. Technometrics 1993;35(2):149–60.
[2] Hansen MH, Nair VN, Friedman DJ. Monitoring wafer map data from integrated circuit fabrication processes for spatially clustered defects. Technometrics 1997;39(3):241–53.
[3] Zhao X, Cui L. Defect pattern recognition on nano/micro integrated circuits wafer. In: NEMS. IEEE; 2008. p. 519–23.
[4] Kainer KU. Metal matrix composites. John Wiley & Sons; 2006.
[5] Chang CJ, Xu L, Huang Q, Shi J. Quantitative characterization and modeling strategy of nanoparticle dispersion in polymer composites. IIE Trans 2012;44(7):523–33.
[6] Zhou Q, Zhou J, De Cicco M, Zhou S, Li X. Detecting 3D spatial clustering of particles in nanocomposites based on cross-sectional images, Technometrics; 2013 [in press], http://dx.doi.org/10.1080/00401706.2013.804440.
[7] Li J, Shi J, Chang T-S. On-line seam detection in rolling processes using snake projection and discrete wavelet transform. J Manuf Sci Eng 2007;129(5):926–33.
[8] Pan E, Shi J, Chang TS, Ye L. On-line bleeds detection in continuous casting processes using engineering-driven rule-based algorithm. J Manuf Sci Eng 2009;131(6):061008-1–8-9.
[9] Schuh A, Woodall W, Camelio J. The effect of aggregating data when monitoring a poisson process. J Quality Technol; 2013 [in press].
[10] Karr AF. Point processes and their statistical inference, 2nd ed., vol. 7. LLC: CRC Press; 1991.
[11] Cowling A, Hall P, Phillips MJ. Bootstrap confidence regions for the intensity of a poisson point process. J Am Stat Assoc 1996;91(436):1516–24.
[12] Tian L, Zucker D, Wei L. On the cox model with time-varying regression coefficients. J Am Stat Assoc 2005;100(469):172–83.
[13] Mohler GO, Short MB, Brantingham PJ, Schoenberg FP, Tita GE. Self-exciting point process modeling of crime. J Am Stat Assoc 2011;106(493):100–8.
[14] Diggle P. A kernel method for smoothing point process data. Appl Stat 1985;34:138–47.