1995

Characterization of surface profiles using discrete measurement systems

Tai-Hung Yang
Iowa State University

Characterization of surface profiles using discrete measurement systems

by

Tai-Hung Yang

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department: Industrial and Manufacturing Systems Engineering
Major: Industrial Engineering

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa
1995

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to express my sincere thanks and appreciation to many people who have encouraged, supported me in my efforts to complete this dissertation.

First, my deepest gratitude is expressed to my major professor Dr. John Jackman. Because of Dr. Jackman's guidance, encouragement and numerous suggestions, it was possible for me to complete this project. His support and expert advise are greatly appreciated.

My appreciation is also extended to the other committee members, Dr. Stephen Vardeman, Dr. John Stufken, Dr. Doug Gemmill, and Dr. Doug McBeth for their valuable contribution to this research.

I also owe my sincerest thanks to my parents Yao-Jen and Meng-Ya for their unconditional love, support, and patience during my studies at Iowa State University.

Finally, a most hearty thanks to my wife and son, Hui-Jane and Stephen, for their support, encouragement, and love they shared with me while completing this dissertation.

CHAPTER 1. INTRODUCTION

Inspection systems that provide metrology information for discrete points on the surface of an object must use some type of fitting procedure to obtain more meaningful dimensional and form information. In this context, we can examine the deviations of the surface being measured and neglect for the moment uncertainties introduced by the inspection system.

Consider the idealized geometric boundary shown in Figure 1.1 where the desired form feature is a straight line and the profile S is the measured profile which is a convolution of the true surface profile and the stylus radius of a coordinate measurement machine (CMM). Form tolerances (flatness, straightness, circularity, and cylindricity) (ANSI Y14.5M, 1982) specify a "zone" within which the toleranced profile or surface must fit. The zone is bounded by two perfect offset profiles or surfaces. We need only to specify the offset value and no datum is needed, i.e., the tolerance zone floats in space. The forthcoming ASME Y14.5.1M-1993 standard gives the mathematical definitions of these geometric tolerances (for details see Walker and Srinivasan, 1993), e.g.,

> **Flatness.** Flatness is the condition of a surface having all elements in one plane. A flatness tolerance specifies a tolerance zone defined by two parallel planes within which the surface must lie.

Definition: A flatness tolerance specifies that all points of the surface must lie in some zone bounded by two parallel planes which are separated by the specified tolerance.

A flatness zone is a volume consisting of all points $\vec{P}$ satisfying the condition

$$|\hat{T} \cdot (\vec{P} - \vec{A})| \leq t/2$$

where:

$\hat{T}$ is the unit direction vector of the parallel planes defining the flatness zone;

$\vec{A}$ is a position vector locating the mid-plane of the flatness zone;

t is the size of the flatness zone (the separation of the parallel planes).

Conformance: A feature conforms to a flatness tolerances t_0 if all points of the feature lie within some flatness zone as defined above, with $t = t_0$. That is, there exist $\hat{T}$ and $\vec{A}$ such that with $t = t_0$, all points of the feature are within the flatness zone.

Actual value: The actual value of flatness for a surface is the smallest flatness tolerance to which the surface will conform.

The other definitions of form tolerances are analogs of the above. In Figure 1.1 the profile S is bounded by a zone that we compare with a specified geometric tolerance zone, t, (i.e., straightness) from the design specifications. For a discrete set of points, we want to determine if S lies within the specified tolerance zone. We are faced with the problem of making inferences about the limits of the zone with incomplete information on S. The limits of S, $l1$ and $l2$, are separated by a distance, w, which must be compared with the tolerance zone specification. In practice, the

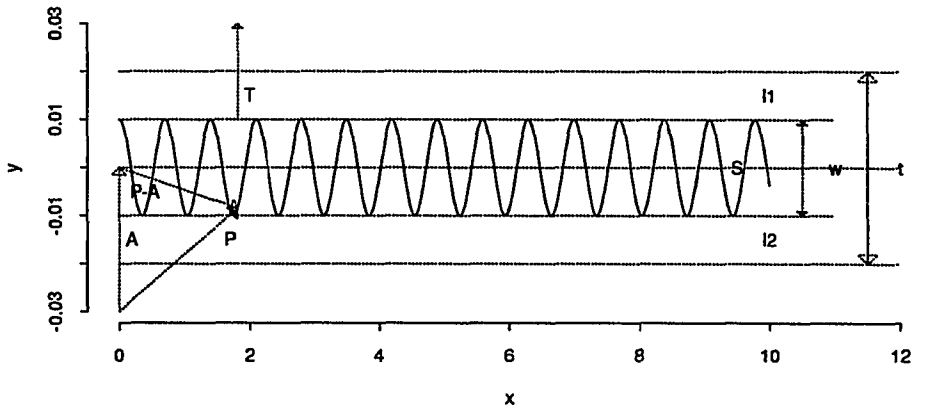


Figure 1.1: Idealized geometric boundary ($y = .01\,cos(.1 - 9\,x)$)

profile S and the supporting lines ($l1$ and $l2$) are unknown because it is infeasible to take an infinite number of points on S. Thus, the sampling strategy (location and number of points) and sample data analysis become critical issues in the context of CMM inspection.

The profile S is usually referred to as one realization of "spatial random processes". The variations occur along the length of the manufactured part. We model these variations as nonstationary harmonic processes, which is not without precedent (Nayak, 1971; Sayles and Thomas, 1978). The surface profiles can be represented as a composite of sine and cosine waves with different amplitudes, phases, and frequencies. This harmonic process model is defined by

$$Z(\mathbf{s}) = \sum_{i=1}^{K} C_i \;\; cos(x\; kx_i + y\; ky_i + \phi_i), \tag{1.1}$$

where $\mathbf{s} = (x, y)$, and there are $K < \infty$ wave-vectors (kx_i, ky_i), $\{C_i\}$ are constants, and the $\{\phi_i\}$, $(i = 1, ..., K)$, are independent random variables which are fixed for a specific surface. The resulting surface profiles are bandlimited functions, i.e., they can

be represented by a finite limit (truncated) inverse Fourier transform (Jerri, 1977).

This research reviews and evaluates the current sampling strategies and sample data analysis for form error evaluation. Without considering the correlation of the sample points, we model the probability distributions for the deviations of the surface profiles and derive the theoretical distribution of form error estimation for random sampling. Further, we consider the spatial dependence between the sample points by using spatial statistics, especially the universal kriging method, to predict form errors. Spline approximation provides optimal theoretical solutions to the estimation of functions from limited data (Powell, 1981). We show that Shannon sampling functions (infinite degree spline interpolation functions) are the optimal approximation of the measured surface profiles from limited sample points. Finally, we use 3D measurement data from actual machined surfaces of common manufacturing processes (Stout et al., 1990) to evaluate form error estimation methods. The performances of universal kriging and a Shannon sampling method on the estimation of form errors are also evaluated for these machined surfaces.

The contribution of this research is to provide a scientific basis for the effect of surface profile, sampling methods, and sample size on the result of form error estimation. This helps to explain flaws in current CMM sampling practice, which can lead to significant errors in form evaluation. We also provide an optimal spline interpolation method for reconstructing surfaces from limited sample points. The performances of this method and kriging are also compared to provide a guideline for limited sample data analysis.

The organization of this dissertation is as follows:

In Chapter 2 the current CMM sampling strategy and sample data analysis are

reviewed. Chapter 3 derives the theoretical distribution of form errors and shows the limitations of discrete sampling. Chapter 4 considers spatial statistics for form error estimation. Chapter 5 discusses the optimal spline interpolation methods. Chapter 6 describes the validation of our models and performance comparisons of an optimal spline method and kriging. Chapter 7 presents conclusions and recommendations for further research in related topics.

CHAPTER 2. LITERATURE REVIEW

Sample data analysis

Suppose we take 19 points uniformly distributed with respect to the x axis. Using standard Monte Carlo techniques, we obtain the points shown in Table 2.1 and Figure 2.1. Researchers have investigated a number of different methods for analyzing sample points such as these. The least squares method is implemented in most CMM software. Murthy and Abdin (1980) and Shunmugam (1986, 1987a, 1987b, 1990, 1991) have already demonstrated that estimates of w obtained from the least squares method do not agree with the definition of form errors. For example, using our points in Table 2.1, the least squares method (shown in Figure 2.2) gives us a value for $\hat{w}$ of 0.0283, which is greater than the true value of 0.02 by 41.5%. Hence, many attempts have been made to derive the minimum values of form errors. The two main approaches to solve this problem have been curve fitting and computational geometry.

Curve fitting approach

Many fitting criteria can be expressed as special cases of a general criterion called Lp-norm estimation. The objective is to find the fit parameters that minimize the

Table 2.1: Nineteen sample points (random sampling)

x	y	orders of y value	Normal distance to the Least squares line	Normal distance to the Mini-max line
0.285092	−0.007802	$x_{(5)}$	−0.01433560*	−0.00782896
0.105460	0.006606	$x_{(12)}$	−0.00014186	0.00657907
3.499036	0.009997	$x_{(18)}$	0.00729671	0.00996947*
3.393897	0.005646	$x_{(10)}$	0.00282031	0.00561849
4.624379	−0.007784	$x_{(6)}$	−0.00914207	−0.00781173
2.800454	0.009996	$x_{(17)}$	0.00646250	0.00996859
9.516238	−0.007497	$x_{(7)}$	−0.00302047	−0.00752561
6.350012	0.008769	$x_{(14)}$	0.00946911	0.00874096
2.904454	0.006158	$x_{(11)}$	0.00274854	0.00613057
8.725649	−0.009941	$x_{(1)}$	−0.00640742	−0.00996947*
1.489229	0.007407	$x_{(13)}$	0.00230958	0.00737983
9.454947	−0.009853	$x_{(2)}$	−0.00544958	−0.00988160
5.553856	0.009284	$x_{(15)}$	0.00903452	0.00925610
9.503680	−0.008196	$x_{(4)}$	−0.00373445	−0.00822461
2.691309	0.005311	$x_{(9)}$	0.00164732	0.00528361
0.007508	0.009995	$x_{(16)}$	0.00313031	0.00996809
9.089049	0.009998	$x_{(19)}$	0.01396500*	0.00996947*
6.554264	−0.006954	$x_{(8)}$	−0.00601026	−0.00698208
4.592405	−0.009246	$x_{(3)}$	−0.01064220	−0.00927373

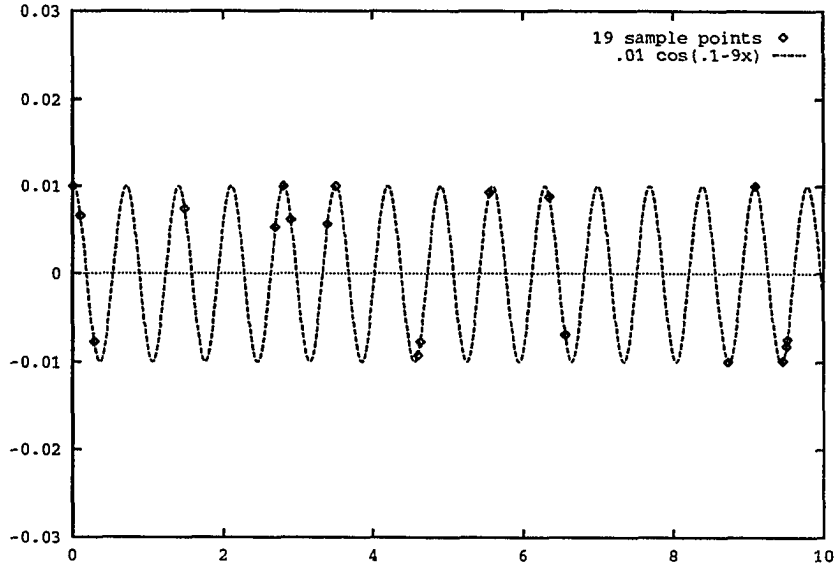


Figure 2.1: Measured sample points (random sampling)

Lp norm (Hopp, 1993)

$$Lp = \left[\frac{1}{n}\sum_{i}|r_i|^p\right]^{1/p} \tag{2.1}$$

where r_i is the i^{th} residual and the sum is over n data points. After the fit parameters are found, the form error is expressed as

$$\hat{w} = |r_{\max}| + |r_{\min}|. \tag{2.2}$$

The residuals for various geometry form deviations are described as follows (Menq et al., 1990, and Shunmugam, 1986, 1987a, 1987b, 1990, 1991).

1. Straightness

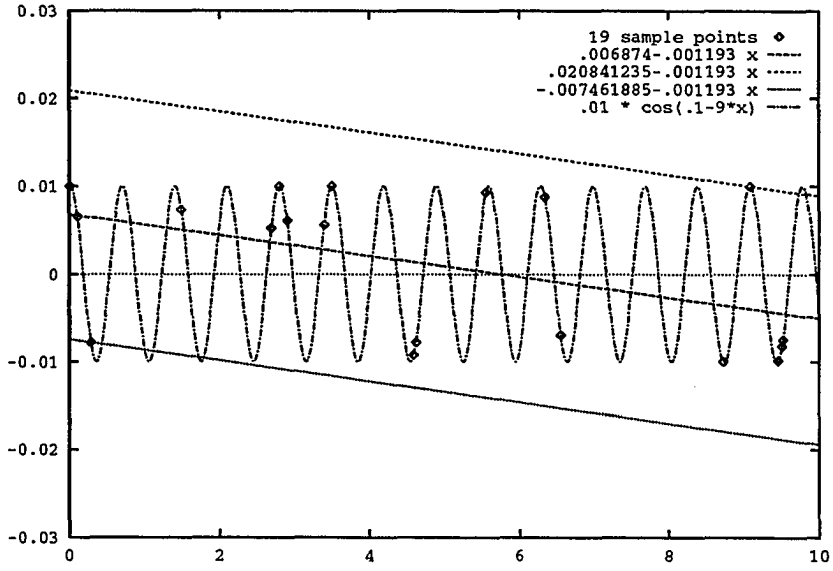


Figure 2.2: Least square fit

The measurement data are given by $\{x_i, y_i\}$ and let the required feature be a straight line of the form

$$y = \eta_1 + \eta_2 x, \tag{2.3}$$

then the normal deviation is expressed as

$$\hat{r}_i = \frac{y_i - (\eta_1 + \eta_2 x_i)}{\sqrt{1 + \eta_2^2}} \tag{2.4}$$

and the linear deviation is

$$\bar{r}_i = y_i - (\eta_1 + \eta_2 x_i). \tag{2.5}$$

2. Flatness

The measurement data are given by $\{x_i, y_i, z_i\}$ and let the required feature be a plane of the form

$$z = \eta_1 + \eta_2\, x + \eta_3\, y, \tag{2.6}$$

then the normal deviation is expressed as

$$\hat{r}_i = \frac{z_i - (\eta_1 + \eta_2\, x_i + \eta_3\, y_i)}{\sqrt{1 + \eta_2^2 + \eta_3^2}} \tag{2.7}$$

and the linear deviation is

$$\bar{r}_i = z_i - (\eta_1 + \eta_2\, x_i + \eta_3\, y_i). \tag{2.8}$$

3. Circularity

The measurement data are given by $\{x_i, y_i\}$ or $\{r_i, \theta_i\}$ and let the required feature be a circle of the form

$$(x - \eta_1)^2 + (y - \eta_2)^2 = \eta_3^2, \tag{2.9}$$

where the center $\{\eta_1, \eta_2\}$ and radius η_3 have to be determined, then the normal deviation is expressed as

$$\hat{r}_i = \sqrt{(x_i - \eta_1)^2 + (y_i - \eta_2)^2} - \eta_3. \tag{2.10}$$

4. Cylindricity

The measurement data are given by $\{r_i, \theta_i, z_i\}$ or $\{x_i, y_i, z_i\}$. The direction vector of the center axis $(\eta_1, \eta_2, 1.)$ and its base point $(\eta_3, \eta_4, 0)$, of the assessment cylinder and the radius η_5 have to be determined. The normal deviation is given by

$$\hat{r}_i = \sqrt{\frac{(x_i - \eta_3 - z_i\eta_1)^2 + (z_i\eta_2 - y_i + \eta_4)^2 + [(y_i - \eta_4)\eta_1 - (x_i - \eta_3)\eta_2]}{\eta_1^2 + \eta_2^2 + 1}} - \eta_5 \tag{2.11}$$

For a cylindrical feature aligned properly with z-axis, the deviation can be expressed in the linear form

$$\bar{r_i} = r_i - [\eta_5 + (\eta_1 + \eta_2\, z_i)\, cos\theta_i + (\eta_3 + \eta_4\, z_i)\, sin\theta_i]. \tag{2.12}$$

Most metrologists only concentrate on the Gaussian (L_2) and the Tschebycheff (L_∞) methodologies (Hocken et al., 1993).

- L_2-estimation

The least squares method minimizes

$$L_2 = \sum \bar{r}_i^2 \tag{2.13}$$

and the normal least squares method minimizes

$$L_2 = \sum \hat{r}_i^2. \tag{2.14}$$

The solutions of (2.14) are discussed in Murthy and Abdin (1980). For straightness, there are two perpendicular lines satisfying the normal least squares fit. The slopes of these two lines are

$$a_1 = \frac{-a_{xy} + \sqrt{a_{xy}^2 + 4}}{2};\ \ a_2 = \frac{-a_{xy} - \sqrt{a_{xy}^2 + 4}}{2}, \tag{2.15}$$

and the intercepts are

$$c_1 = \frac{\sum(y_i - a_1\, x_i)}{N};\ \ c_2 = \frac{\sum(y_i - a_2\, x_i)}{N}, \tag{2.16}$$

where $a_{xy} = \frac{[N\sum x_i^2 - (\sum x_i)^2] - [N\sum y_i^2 - (\sum y_i)^2]}{[N\sum x_i y_i - \sum x_i \sum y_i]}$. Obviously, only one of them is correct for the best fit. For flatness and circularity, the solution equations are complex and cannot be solved easily. They use trial and error procedures to solve this problem.

- L_∞-estimation

  The limit of L_p as p goes to infinity is the largest magnitude residual, so the L_∞ problem is to minimize the maximum magnitude residual, i.e.,

$$\min \max |\hat{r}_i|. \tag{2.17}$$

  Shunmugam (1986, 1987a, 1987b, 1990), Dhanish and Shunmugam (1991), and later Caskey et al. (1992) all try to develop various search procedures to solve this problem. Gonin and Money (1989) investigate the current available algorithms and FORTRAN programs for solving this type of problem. We use the Anderson-Watson-Osborne algorithm and program suggested by Gonin and Money to solve (2.17).

Computational geometry approach

The width of a set of measurement points P in two dimensions is the minimum distance between parallel lines of support of P, which defines the straightness error. In three dimensions, it is the minimum distance between parallel planes of support, which defines the flatness error. Houle and Toussaint (1988) present algorithms to calculate the width of a set of points in two and three dimensions, i.e., the minimum width of the convex hull. Etesami and Qiao (1989) and Traband et al. (1989) use the same approach to evaluate the straightness and flatness error. For the circular features, the determination of two concentric circles with minimum distance enclosing the measurement points requires the application of Voronoi diagrams, e.g., Etesami and Qiao (1989) and Roy and Zhang (1992) present this approach to evaluate circularity.

Houle and Toussaint (1988) also bring up an open problem concerning the related area of minimax approximation. The width of a set they presented solves the problem of finding a minimax approximating line of a set of points in the plane as well as that of finding the minimax approximating plane of a set of points in three dimensions. That is, they have proved that the minimum width of the convex hull of the measurement points is equal to the distance to the minimax-approximating-line (or plane). However, these only solve the straightness and flatness errors. For other general type of geometric errors, e.g., finding the minimax approximating line of a set of points in three dimensions, the computational geometry approach is still an open research problem. Caskey et al. (1992) reach the same conclusion that the "minimum zone is the same as mini-max (L_∞)". That is, the value of $\hat{w}$ after solving (2.17) will have the same value as the width of a set determined by the minimum width of its convex hull.

Application

We implement a set of programs, both L_p-norm estimation and width of convex hull, to calculate the form errors (straightness, flatness, circularity, and cylindricity). Appendix A lists these programs and their algorithms. Appendix B presents the computation results of published data sets and gives the comparisons of the published results in the literature.

Questions remain to be answered

For our nineteen sample points example, we obtain an estimate of w of 0.019939 by either solving (2.17) or calculating the width of its convex hull, which is shown in

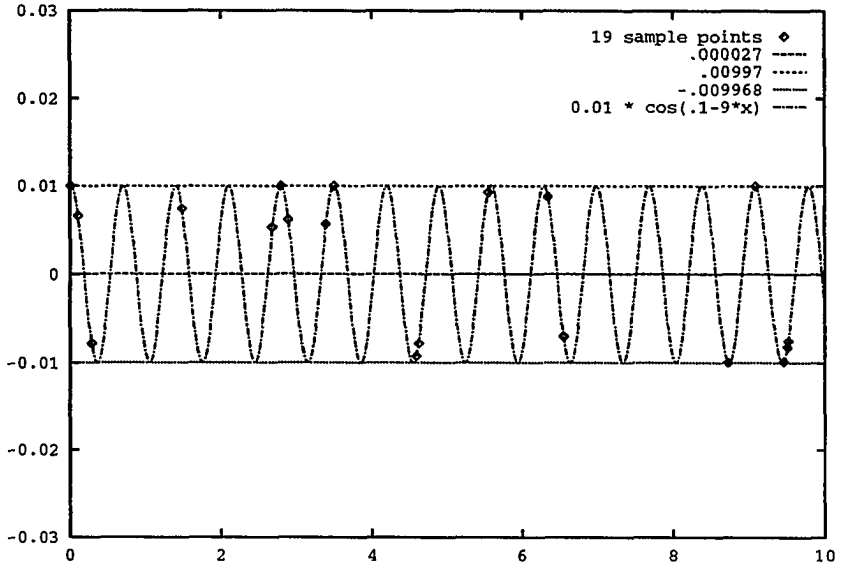


Figure 2.3: Minimax method

Figure 2.3. Therefore, we might conclude that this is a good sampling strategy and analysis, because we are very close to the true value.

However, from Figures 1.1 to 2.3, we recognize that $\hat{w}$ is only an estimate of the true value w. As Menq, et al. (1990) argue that "cases may occur when all sampled deviations are found to be within tolerance while some non-sampled deviations are in fact out of tolerance. Therefore, using CMM sampling for inspection does not meet the principle of minimum zone method." Based on the geometric tolerance definition, the minimum zone principle is valid only under the assumption that the entire profile is examined.

Further, Hopp (1993) considers the general metric p of L_p-norm and reports that the bias and sensitivity errors of the fit will vary with p. As p increases, the

sensitivity of the fit to point measurement error increases, but the bias with respect to the fit prescribed by tolerance theory decreases. Therefore, it is very difficult to develop general guidelines for the proper choice of fitting objective (p value) for a practical coordinate measuring system. Least-squares fitting is widely used and debated, with many claiming that extremal fitting is better because it "conforms" to tolerance theory. An alternative fitting objective is to minimize combined uncertainty in the result. How to make the best choice is not at all clear (Hopp, 1993). Thus, the questions remain, how good is the estimate? How will different values for the p metric of the L_p-norm estimation affect this estimate?

Sampling strategies

Sampling strategies involve two issues: determination of locations and number of sample points needed to inspect a workpiece. Locations of sample points can be generated by a number of different sampling techniques such as uniform sampling, random sampling (the example we used in previous section), stratified random sampling, Hammersley sampling, etc. (Hocken et al., 1993, Caskey et al., 1991, 1992, and Woo et al., 1993). Hocken, et al. (1993) point out that uniform sampling is the most practical for users since it is much easier to program a measuring machine for equal intervals in angle or space than it is to use any other sampling techniques. Caskey, et al. (1992) prefer stratified random sampling and argue that it is more robust to feature waviness based on their simulation results. Even with this method, their results gave poor performance in estimation of form errors. They found that a large number of samples is necessary to obtain a low variance of the estimated form error based on a particular surface. The relationship between feature waviness and

the sampling methods is still unknown.

The number of required measurement points has very important implications on inspection speed and data analysis. Ideally, we would like an inspection plan that requires the fewest measurements yet provides sufficient information for quality assurance. However, the number of required measurement points for obtaining effective quality assurance is unknown, and its determination also still remains unaddressed (Menq et al., 1990).

We begin our discussion with a probabilistic view of form deviations. Using known surfaces, we show how random sampling methods and sample size affect form error estimation. We conclude with a comparison of analysis techniques and a discussion of the role of the metric p in L_p-norm estimation. We will also discuss the other fitting procedures, such as kriging and spline interpolation, which consider the spatial correlations between points, whether they have better estimate of form errors or not.

CHAPTER 3. STATISTICAL DISTRIBUTIONS FOR SURFACE PROFILES AND ITS IMPACT ON SAMPLING RESULTS

Definition 1: L_p mean profile level: L_p mean profile level defines a profile that minimizes (2.1). For $n \to \infty$ and all measurement points (2D or 3D) within the domain of inspection, the mean profile level is defined as the *true* L_p mean profile level, which is unknown to us. For n inspection points, the mean profile level is defined as the *measured or estimated* L_p mean profile level.

Definition 2: Minimum zone mean profile level: the L_∞ mean profile level is defined as the minimum zone mean profile level. As we discussed in the previous chapter, the L_∞ mean profile level minimizes the maximum residual, which is the width of the minimum zone and is the form error as defined in the standard. The statistical distribution we will discuss is distributed along this minimum zone (L_∞) mean profile level.

Definition 3: Detectability of true form error. Detectability is defined as the ratio of the estimated form error for n sample points to the true form error. For example, for $n = 3$, we estimate $\hat{w} = 0.012$ for $w = 0.02$. Therefore, the detectability is 60% (0.012/0.02) to detect the true form error with $n = 3$. We will show next that the variance of $\hat{w}$ is monotonically decreasing with n and thus the mean is enough to define the detectability.

Deviation distribution from L_∞ mean profile level

The periodic surface profile shown in Figure 1.1 can be represented by the general sine wave function

$$y(x) = A\,Cos(bx + c), \tag{3.1}$$

where, A is the amplitude, b is the frequency and c is the phase angle. Now suppose we perform random sampling (uniformly) along x axis on the interval $(0, x_0)$, that is, x has a uniform probability density function (pdf) $f(x)$ given by

$$f(x) = \begin{cases} \frac{1}{x_0} & 0 \le x \le x_0 \\ 0 & otherwise. \end{cases}$$

We would like to determine the pdf of y. For one value of y, two possible values of x will match it within one cycle. There are a total of $\frac{x_0 b}{2\pi}$ cycles within $x \in (0, x_0)$. Thus, the inverse function $y^{-1}(x)$ is a real $\frac{x_0 b}{\pi}$-valued function of y, where $\frac{x_0 b}{\pi}$ is forced to be an integer and all of the $\frac{x_0 b}{\pi}$ values have equal probability. Hence,

$$\begin{aligned} f(y) = & \frac{Pr\{y<y(x)\le y+\Delta y\}}{\Delta y} \\ & = \frac{x_0 b}{\pi}\frac{Pr\{x<x(y)\le x+\Delta x\}}{\Delta y} \\ & = \frac{x_0 b}{\pi}\frac{Pr\{x<x(y)\le x+\Delta x\}}{\Delta x}\frac{\Delta x}{\Delta y} \\ & = \frac{x_0 b}{\pi} f(x)\left|\frac{dx}{dy}\right| \\ & = \frac{b/\pi}{|dy/dx|} \quad \text{for} \quad \frac{dy}{dx} \neq 0 \end{aligned}$$

where

$$\left|\frac{dy}{dx}\right| = AbSin(bx + c) = Ab\sqrt{1 - Cos^2(bx + c)} = b\sqrt{A^2 - y^2}.$$

Thus,

$$f(y) = \begin{cases} \frac{1}{\pi\sqrt{A^2 - y^2}} & |y| < A \\ 0 & |y| \ge A. \end{cases} \tag{3.2}$$

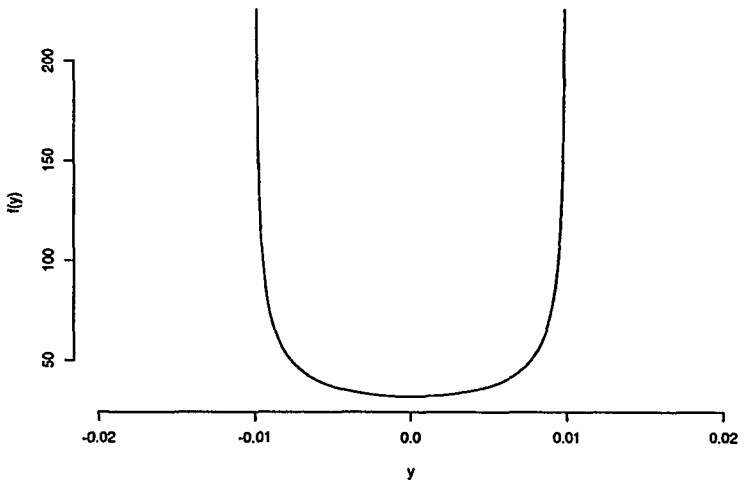


Figure 3.1: $f(y) = \frac{1}{\pi(0.01^2 - y^2)^{0.5}}$ for $|y| < 0.01$; 0 for $|y| \geq 0.01$

This probability density function for profile S is shown in Figure 3.1.

From (3.2) and Figure 3.1 we would like to emphasize the following points.

1. This distribution is quite different from the normal distribution. The tails of the distribution play an important role in evaluating the form errors as we will show in the next section.

2. The pdf in (3.2) is not a function of frequency. Thus, detectability using random sampling is the same under different frequencies with this type of periodic profile.

3. Since the domain of this pdf is the amplitude A, the distribution is the same with different amplitudes of $y(x)$. Thus, the detectability using random sampling is the same irrespective of different amplitudes of $y(x)$.

The minimum zone mean profile levels are difficult to obtain and the derivation of the sine wave distribution in (3.2) becomes more cumbersome when we consider more complicated profiles, e.g., profile B: $y(x) = 0.25(sin\ 3x + cos\ 12x + sin\ .5x + cos\ 5x)$ for $x \in (0, 50)$, which is shown in Figure 3.2. An alternative approach is to approximate the minimum zone mean profile levels and the statistical distribution numerically by generating a sequence of points for fixed increments of x, running a mini-max optimization program to find the parameters, and creating a frequency histogram from the residual values. For example, we generate points for the interval 0.01 between points along x axis ($x \in [0, 50]$). We have 5001 points which are used to approximate the continuous profile B. Then, we obtain the L_2 (least squares) mean profile level as $-0.002868x + 0.071354$ and the minimum zone (minimax) mean profile level as $-0.00001x + 0.047791$, for $x \in [0, 50]$, which are shown in Figure 3.3. Thus, the deviations from the minimum zone (L_∞) mean profile level are no longer $y(x) = 0$. Figure 3.4 shows the height (normal distances to the minimum zone mean profile level) distribution function for profile B.

From these two examples of surface profiles, it is clear that the normal distribution assumption to estimate the distribution of CMM measured surface profile that result from manufacturing processes is not always applicable. Sayles and Thomas (1978) gave strong support of this argument.

> With a surface of a given finite area, if the heights are measured sufficiently closely together then an arbitrarily large sample can be obtained. Its extreme values, however, would not be very large and will therefore be a function not of the sample size, as for a discrete normal variate, but of the area over which the sample was taken. This is because in the physical

> world we do not expect things to vary by vast amounts in short periods or distances. Large changes in surface height on Gaussian surfaces are possible but they tend to occur over large distances. This argument suggests that in physical situations a Gaussian stationary process is in fact a contradiction in terms. (p. 432)

Stout et al. (1990) has shown the height distributions for various machined surfaces. Many distributions shown in that book cannot be modeled as normal distributions. We use the beta distribution, which is not without precedent (He, 1991), to estimate the height distribution function, i.e., the form deviation from the L_∞ mean profile level, for the profiles measured by CMM. By changing the parameters of the beta distribution, we can control its shape to match that of a specific process. The probability density function for a generalized beta distribution is given by

$$f(X,\alpha,\beta,a,b) = \frac{1}{(b-a)B(\alpha,\beta)}\left(\frac{X-a}{b-a}\right)^{\alpha-1}\left(1-\frac{X-a}{b-a}\right)^{\beta-1},$$

$$\alpha > 0, \beta > 0, a \leq X \leq b,$$

$$\text{where, } B(\alpha,\beta) = \int_0^1 Z^{\alpha-1}(1-Z)^{\beta-1}dZ,$$

X = random variate (form deviation), a, b = lower and upper limits of the distribution, and α, β = shape exponents. Outside the interval $[a, b]$ the probability density is zero.

We can represent the density function of (3.2) as a beta distribution with $\alpha = \beta = 0.5$, $a = -A$, and $b = A$. It can be seen that the true form (straightness) error for profile B is 1.828712. The histogram shown in Figure 3.4 can be approximated by a beta distribution with $\alpha = 2.68287$, $\beta = 2.97903$, $a = -0.914587$, and $b = 0.914485$, using He's algorithm (1991).

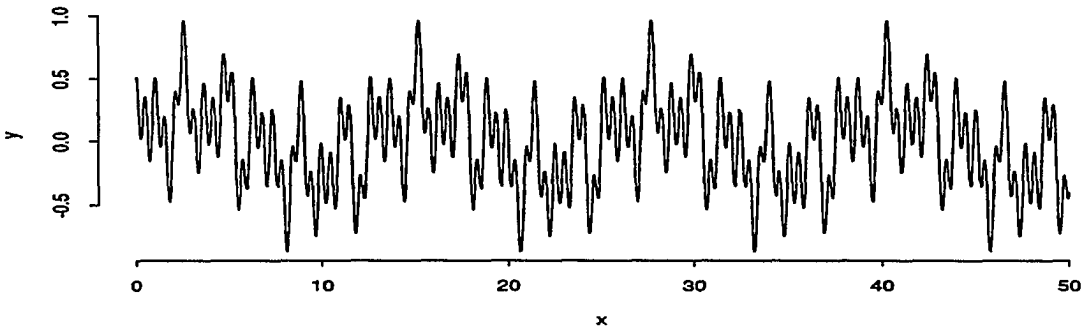


Figure 3.2: Profile B: $y(x) = 0.25(sin\ 3x + cos\ 12x + sin\ .5x + cos\ 5x)$ for $x \in (0, 50)$

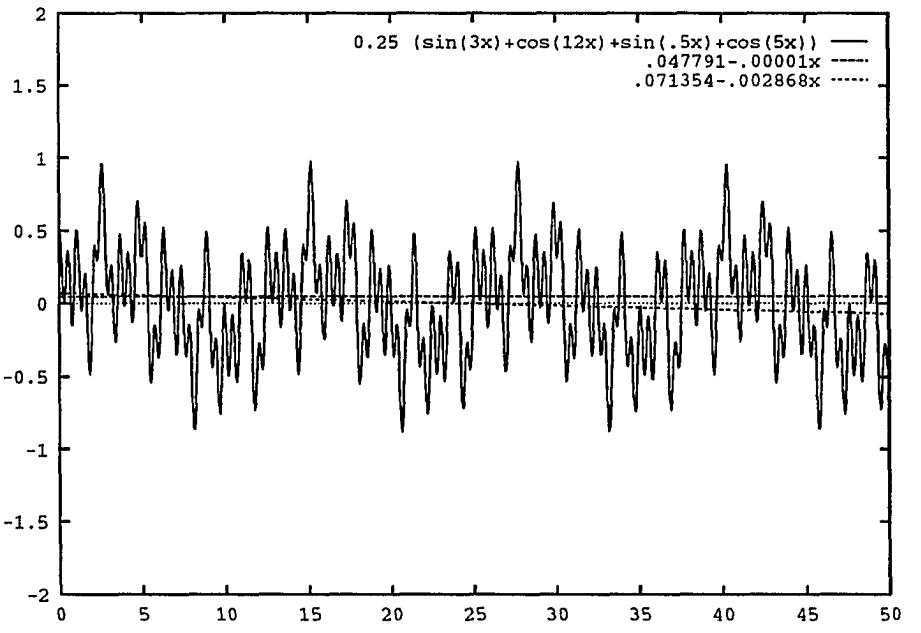


Figure 3.3: L_2 and L_∞ mean profile levels for profile B

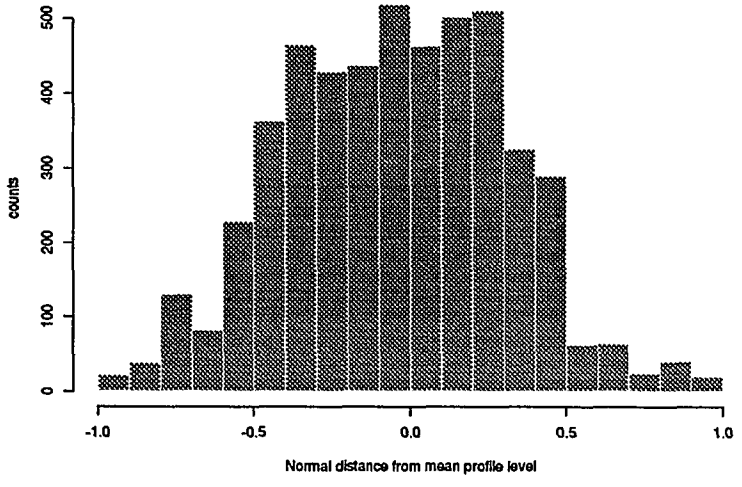


Figure 3.4: Height distribution function for profile B: $y(x) = 0.25(sin\ 3x + cos\ 12x + sin\ .5x + cos\ 5x)$ for $x \in (0, 50)$

Estimated form error and order statistics

Let the minimum zone mean profile level as $y(x) = 0$ and $y_1, y_2, ..., y_n$ be the sample points taken from this inspected profile and $y_{(1)}, y_{(2)}, ..., y_{(n)}$ are the points sorted in ascending order of magnitude such that

$$y_{(1)} \leq y_{(2)} \leq \cdots \leq y_{(n)}.$$

If the minimum zone mean profile level is known, then the samples $y_{(1)}$ and $y_{(n)}$ determine the estimated form error as $y_{(n)} - y_{(1)}$. Since we take random samples (following a uniform distribution) from the inspected profile along the x axis (and if the samples are far enough from each other), we can treat $Y_1, Y_2, ..., Y_n$ as i.i.d. continuous random variables with common beta distribution with parameters $(\alpha, \beta,$

a, b). Let R be the difference (the estimated form error) given by $Y_{(n)} - Y_{(1)}$, then the density function of R (Kendall and Stuart, 1977) for the general case is

$$g_R(r) = n(n-1)\int_{-\infty}^{\infty}\{F(y_{(1)}+r) - F(y_{(1)})\}^{n-2}\ f(y_{(1)})f(y_{(1)}+r)dy_{(1)}. \quad (3.3)$$

Its expected value is

$$E(R) = \int_{-\infty}^{\infty}\{1 - F(y)^n - [1-F(y)]^n\}dy \quad (3.4)$$

and the mth moment about the mean is given by

$$E(R-E(R))^m = m(m-1)\int_{-\infty}^{\infty}\int_{-\infty}^{y_{(n)}}\{1 - F_n^n - (1-F_1)^n + (F_n - F_1)^n\}$$

$$\{R-E(R)\}^{m-2}dy_{(1)}dy_{(n)} - (m-1)\{-E(R)\}^m \text{ for } m \geq 2. \quad (3.5)$$

From the definition of variance we obtain

$$Var(R) = E(R-E(R))^2 =$$

$$2\int_{-\infty}^{\infty}\int_{-\infty}^{y_{(n)}}\{1 - F_n^n - (1-F_1)^n + (F_n - F_1)^n\}dy_{(1)}dy_{(n)} - E(R)^2. \quad (3.6)$$

Using these statistics of the form deviation estimate, we can calculate the detectability with respect to n for given parameters of any distribution. Figure 3.5(a)(b) are the results calculated from (3.4) and (3.6) for different α, β parameters of beta distributions with $a = 0$ and $b = 1$. From these figures we would like to illustrate the following points.

1. These figures assume that the minimum zone mean profile level is known, and the sample points are i.i.d. beta distribution when we perform random sampling.

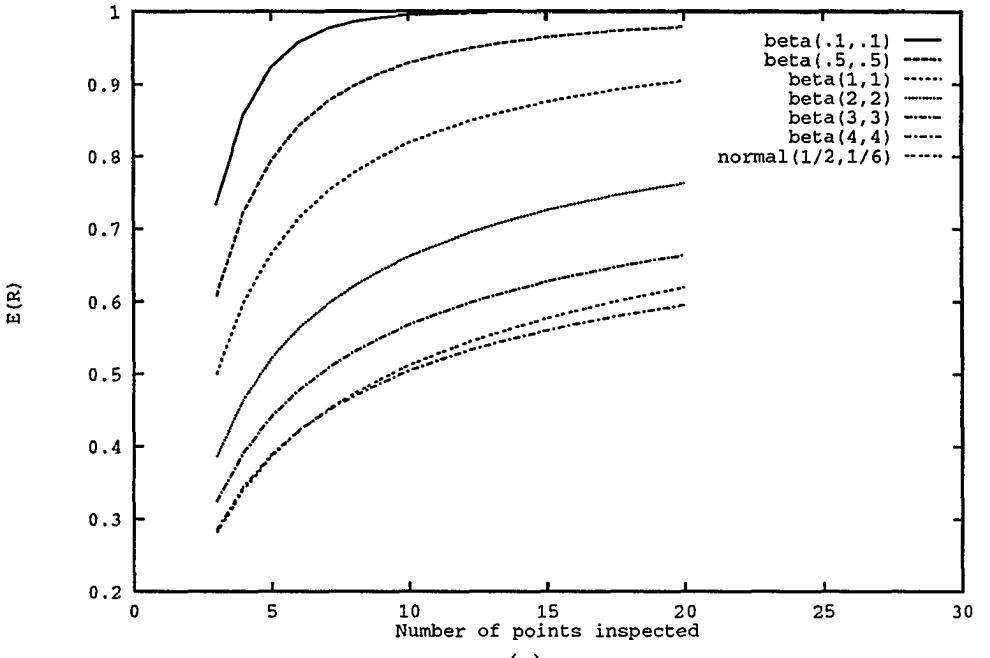
2. As sample size increases, detectability increases and the standard deviation decreases.

3. The detectability for the normal distribution $N(1/2, 1/6)$, which is close to a unit beta distribution with $\alpha = \beta = 4$ and is the most common form deviation distribution assumption for most researchers, is only 62% when we take 20 sample points. Note that, in practice, typically less than 10 points are used to assess form errors. Therefore, a small sample size has a very high probability of accepting a bad part if the measured zone is the same order of magnitude as the tolerance specification. The scenario is even worse when the form deviation follows a beta distribution with parameters (4,4) or higher kurtosis. The mean of estimated form error is only 74% of the true value when we take 100 points (82% for the normal distribution, which is close to beta distribution with $\alpha = \beta = 4$).

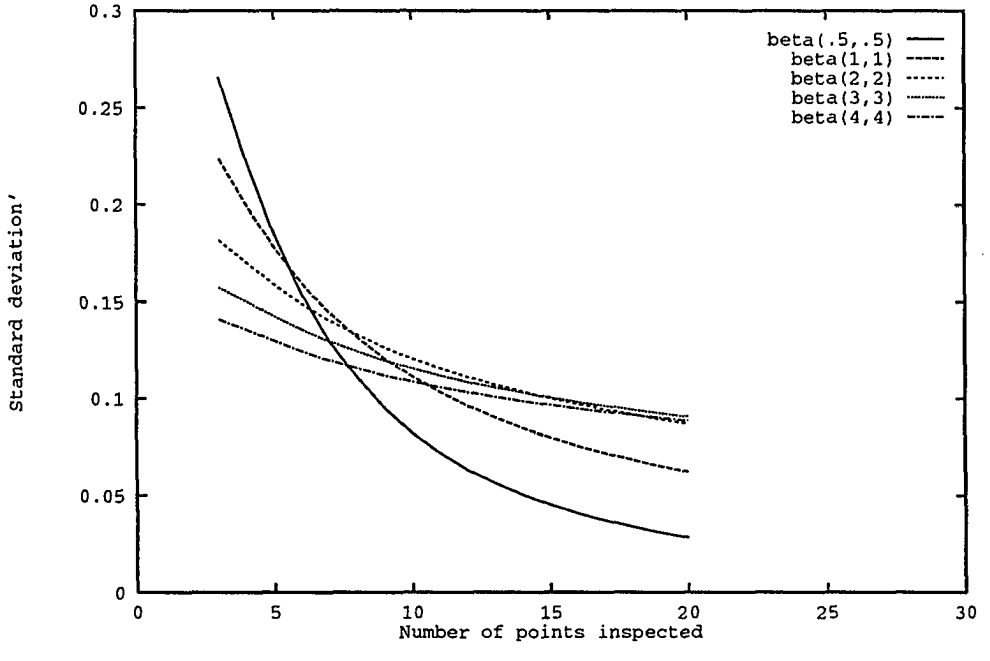
Since the true minimum zone mean profile level is unknown to us, the practice is to take an arbitrary number of sample points to estimate it and then obtain $\hat{w}$. Next, we will use simulation techniques to take sample points from profile S and B and use least squares and mini-max methods to estimate w for different number of sample points which will be compared with the theoretical values.

Simulation results

Varying n from 3 to 20, we generate 30 sets of random points from profile S: $y(x) = 0.5\ Cos(-9x + 0.1)$ for each n and then use the least square method and mini-max method to estimate w. The estimated form error $y_{(n)} - y_{(1)}$ values (given



(a)



(b)

Figure 3.5: (a) Mean (b) Standard deviation of range from beta distributions

a known minimum zone profile level) are also calculated. Figure 3.6 shows the mean values and standard deviations of $\hat{w}$ and the values calculated by (3.4) and (3.6) using a beta distribution with $a = -0.01$, $b = 0.01$ and $\alpha = \beta = 0.5$. For profile B, Figure 3.7 shows the mean values and standard deviations of $\hat{w}$ and the values calculated by (3.4) and (3.6) using a beta distribution with $a = -0.914587$, $b = 0.914485$, $\alpha = 2.68287$, and $\beta = 2.97903$.

Based on these simulation results, we make the following observations:

1. The difference between the mean values calculated from (3.4) and the mean estimate form errors by random sampling from both profiles is getting smaller as the sample size increases. We attribute the bias to the estimate of the minimum zone mean profile level. Small sample sizes exhibit large biases.

2. The mean estimated form error by the mini-max method is smaller than the mean of R calculated from Eq. 3.4 for both profiles. This can result in accepting a part which is actually out of tolerance. Therefore, the type II error (the probability of accepting a bad part) is higher when using the mini-max method.

3. Hopp (1993) suggests that an optimal p value can be found for the parameters of a substitute geometry (i.e., the minimum zone mean profile level), though the way to make the best choice of p is not clear. The mean values $y_{(n)} - y_{(1)}$ we calculated are based on the known minimum zone mean profile level and agree quite well with the mean values calculated from (3.4). We observe that even with good estimates of the minimum zone mean profile level, the detectability is still quite low (e.g., profile B) due to the sampling limitation.

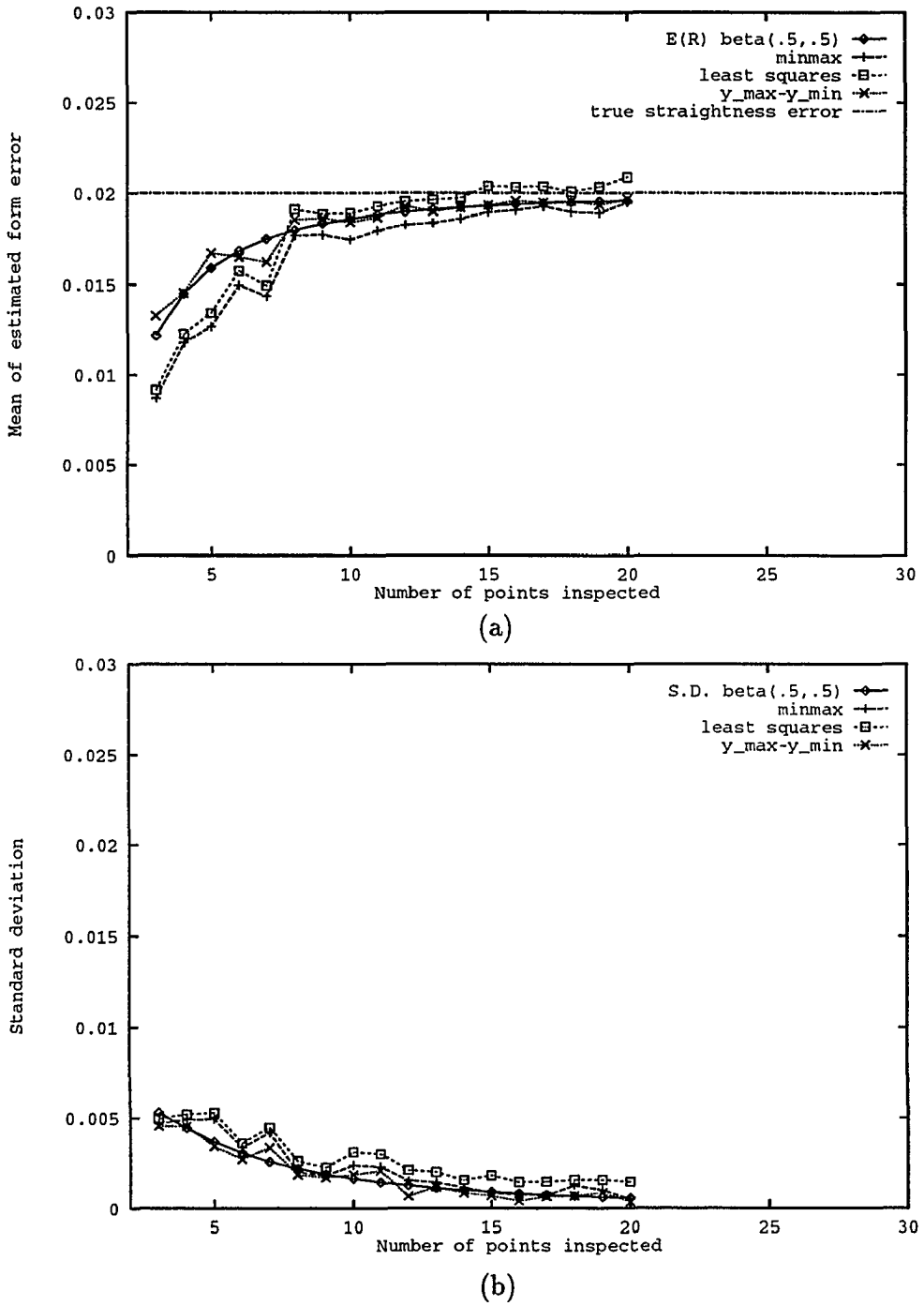


Figure 3.6: (a) Mean comparison (b) Standard deviation comparison of simulation and theoretical results (Profile S)

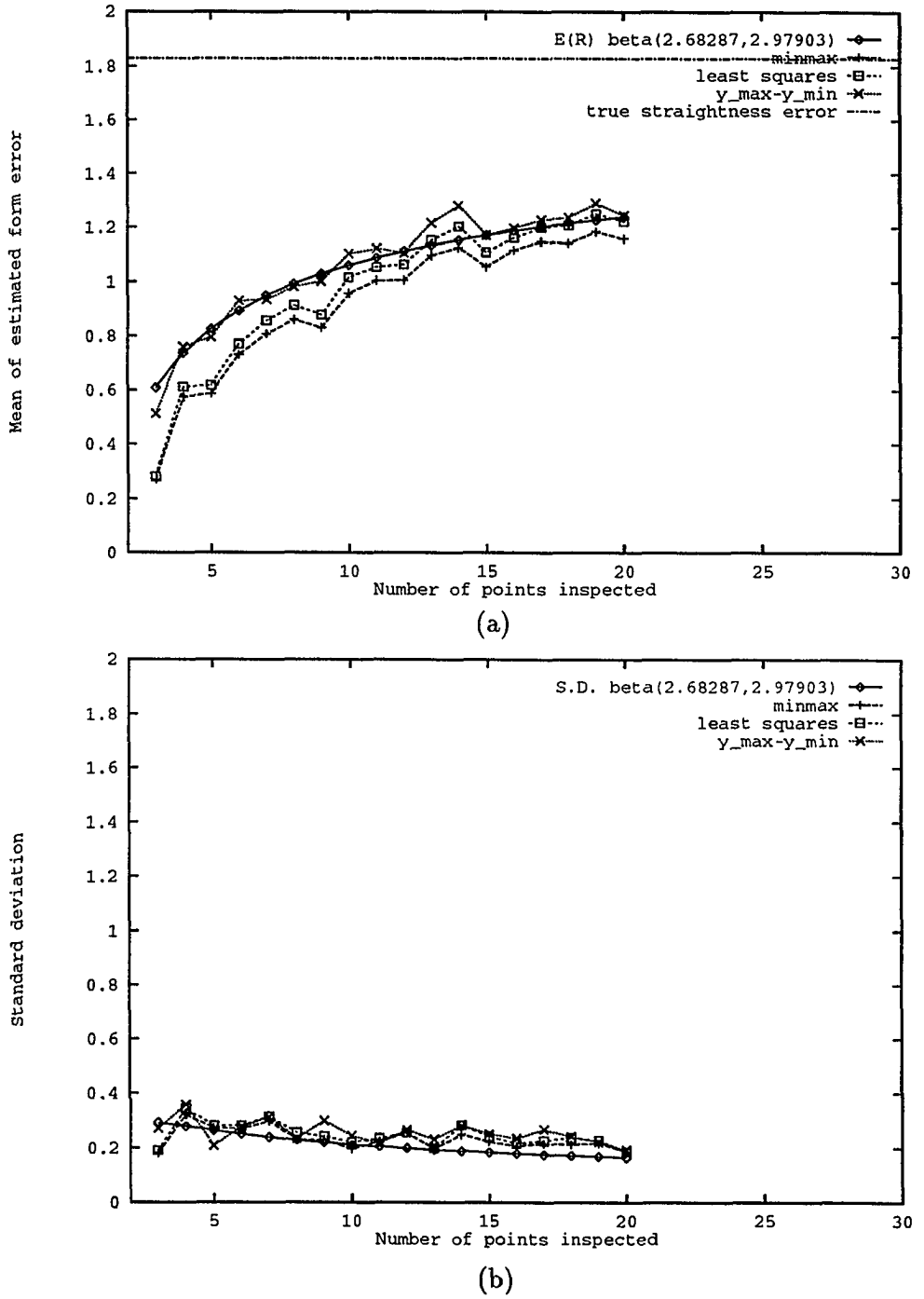


Figure 3.7: (a) Mean comparison (b) Standard deviation comparison of simulation and theoretical results (Profile B)

In order to validate the above observations for a wide variety of surfaces, we conducted another simulation study. We assume $z(x) = 0.5$ as the minimum zone mean profile level, the true straightness error as 1, and the inspection length as 10. Points on the surface are i.i.d. having a beta distribution with parameters (1,1), (1,2), ..., (1,5), (2,1), ..., (2,5), ..., (5,5). Five hundred samples are generated for each distribution with 5 and 50 measurement points for each sample. Mini-max and normal least-squares methods were used to calculate the straightness error for each sample and the $z_{(n)} - z_{(1)}$ value is also calculated.

Table 3.1 and 3.2 list the mean and standard deviation of 500 samples of 5 and 50 measurement points for each sample with various beta distributions. The theoretical mean/s.d.(standard deviation) columns in these tables are calculated from (3.4) and (3.6). The next three columns list the mean and standard deviation for $z_{(n)} - z_{(1)}$ and estimated form errors by using normal least square and mini-max methods. Figure 3.8 plots the mean values in the theoretical mean/sd column of Table 3.1. From this figure we observe most of the mean sample ranges are less than 50% of the true straightness error and they are decreasing as the beta parameters (α, β) are increasing. Figures 3.9(a) and 3.10(a) show mean and standard deviation values of the estimated form errors calculated from $z_{(n)} - z_{(1)}$, normal least-squares and minimax methods for 500 samples with 5 and 50 measurement points/sample respectively. The abscissas of these two figures represent the distribution number and the corresponding beta distributions are shown in Table 3.1 and 3.2. Figures 3.9(b) and 3.10(b) also show the theoretical density function calculated from (3.3) and the estimated form error histograms for $z_{(n)} - z_{(1)}$, mini-max, and normal least-squares methods for beta distribution $(\alpha, \beta) = (2,2)$.

The mean estimated form errors calculated from mini-max consistently have the smallest value. The difference is larger for smaller sample sizes which agrees with our previous observation. The normal least-squares method has a higher probability of over-estimating the form error (greater than 1 in this case) especially when $\alpha \leq 1$ and $\beta \leq 1$. The histogram of estimated form error by using normal least squares (Figure 3.10(b)) shows the estimated errors are greater than 1 for some samples. With the relatively large standard deviation for all cases, we cannot differentiate statistically between the mini-max or normal least-squares estimates.

The inspection length was 10 in the above simulation study. Equations (2.15) and (2.16), the slopes and intercepts of normal least square solution, show that they are not x-axis scale-invariant, i.e., the form error of (x_i, y_i) is different from $(\alpha x_i, y_i)$, where $\alpha > 0$ is the scale. Figure 3.11 also shows the mini-max method is not x-axis scale-invariant. Thus, we conduct another simulation study with an inspection length of 1000 while holding other conditions constant. Table 3.3 and 3.4 show the results. The results are almost identical to the results for an inspection length of 10. From these simulation studies, we observe that the detectability is strongly influenced by the sample size and the surface characteristic. The p metric selection does not seem to affect the detectability significantly.

Table 3.1: Mean and standard deviation for 500 samples, 5 points/sample, inspection length=10

Distribution Number	Beta para. α	β	Theoretical mean/s.d.	500 samples, 5 points/sample $y_{(n)} - y_{(1)}$	L_2	L_∞
1	1	1	0.666667	0.6828517	0.5856701	0.5545228
			0.178174	0.1736909	0.1959300	0.1839747
2	1	2	0.539683	0.5306998	0.4569302	0.4352397
			0.158999	0.1842183	0.1880529	0.1798774
3	1	3	0.436813	0.4310106	0.3718554	0.3534554
			0.165936	0.1620427	0.1610743	0.1536121
4	1	4	0.364002	0.3696611	0.3126132	0.2970302
			0.150774	0.1448102	0.1388524	0.1317659
5	1	5	0.311147	0.3153624	0.2692456	0.2568470
			0.136396	0.1356017	0.1329954	0.1275964
6	2	2	0.521479	0.5244642	0.4485602	0.4262693
			0.158999	0.1603416	0.1622875	0.1537526
7	2	3	0.465182	0.4734989	0.4057428	0.3836184
			0.150519	0.1475860	0.1518935	0.1440550
8	2	4	0.412264	0.4073168	0.3508049	0.3325445
			0.141496	0.1377190	0.1386368	0.1310568
9	2	5	0.367707	0.3584276	0.3038960	0.2893923
			0.132217	0.1251407	0.1251233	0.1190757
10	3	3	0.441251	0.4404719	0.3763416	0.3576119
			0.142519	0.1432202	0.1477351	0.1405761
11	3	4	0.407992	0.4145075	0.3521781	0.3344445
			0.135361	0.1335398	0.1399267	0.1324680
12	3	5	0.375367	0.3835522	0.3334500	0.3158700
			0.128223	0.1335189	0.1340655	0.1271485
13	4	4	0.389125	0.3848140	0.3256564	0.3095435
			0.129743	0.1315119	0.1320908	0.1257106
14	4	5	0.366591	0.3586658	0.3053390	0.2900945
			0.124091	0.1236935	0.1289402	0.1228553
15	5	5	0.351875	0.3518370	0.3030063	0.2870403
			0.119702	0.1168656	0.1130808	0.1062241

Table 3.2: Mean and standard deviation for 500 samples, 50 points/sample, inspection length=10

Distribution Number	Beta para. α	β	Theoretical mean/s.d.	500 samples, 50 points/sample $y_{(n)} - y_{(1)}$	L_2	L_∞
1	1	1	0.960784	0.9627672	0.9896147	0.9484121
			0.0269179	0.02547632	0.05224548	0.03133454
2	1	2	0.865698	0.8700323	0.8845990	0.8556719
			0.0648395	0.06315896	0.07114728	0.06404775
3	1	3	0.752055	0.7521097	0.7632886	0.7394696
			0.0870115	0.08982415	0.09360024	0.09046787
4	1	4	0.655221	0.6633374	0.6699312	0.6531555
			0.0945637	0.09597787	0.09629015	0.09541790
5	1	5	0.577135	0.5785034	0.5834723	0.5690922
			0.0951972	0.09716932	0.09985643	0.09675174
6	2	2	0.85155	0.8582640	0.8599053	0.8322046
			0.0552998	0.05381828	0.06145103	0.05699484
7	2	3	0.78708	0.7866040	0.7871082	0.7614175
			0.0661313	0.06667817	0.07113217	0.06672352
8	2	4	0.717354	0.7131574	0.7130789	0.6914989
			0.0741036	0.07549051	0.07544360	0.07280379
9	2	5	0.65352	0.6500484	0.6490527	0.6290505
			0.077789	0.07668453	0.07813645	0.07470217
10	3	3	0.759056	0.7635570	0.7587376	0.7333470
			0.0649189	0.06410962	0.06997796	0.06470722
11	3	4	0.714093	0.7092233	0.7076002	0.6843604
			0.0680278	0.06971229	0.07141196	0.06827329
12	3	5	0.66669	0.6659320	0.6615880	0.6406171
			0.0704798	0.06886642	0.06958961	0.06646528
13	4	4	0.688286	0.6841785	0.6784570	0.6572393
			0.0674018	0.07032887	0.07170390	0.06787789
14	4	5	0.655064	0.6551808	0.6490667	0.6264009
			0.0679683	0.06779035	0.07012162	0.06447985
15	5	5	0.633213	0.6337910	0.6300958	0.6072051
			0.067193	0.06719966	0.06765132	0.06521175

Table 3.3: Mean and standard deviation for 500 samples, 5 points/sample, inspection length=1000

Distribution Number	Beta para. α	β	500 samples, 50 points/sample $y_{(n)} - y_{(1)}$	L_2	L_∞
1	1	1	0.6655831	0.5786854	0.5474520
			0.1743385	0.1975065	0.1862461
2	1	2	0.5437005	0.4708807	0.4471396
			0.1819644	0.1844359	0.1754473
3	1	3	0.4354061	0.3755207	0.3575147
			0.1661413	0.1621799	0.1559837
4	1	4	0.3537382	0.3065658	0.2919180
			0.1509290	0.1499126	0.1446375
5	1	5	0.3177277	0.2769228	0.2632968
			0.1382845	0.1324432	0.1264102
6	2	2	0.5244169	0.4456181	0.4230248
			0.1599947	0.1680792	0.1600093
7	2	3	0.4678887	0.4094903	0.3881842
			0.1478950	0.1541772	0.1452543
8	2	4	0.4127547	0.3579712	0.3417480
			0.1390793	0.1450233	0.1389728
9	2	5	0.3657815	0.3134359	0.2982966
			0.1255167	0.1255399	0.1199531
10	3	3	0.4417505	0.3745434	0.3555039
			0.1507178	0.1532916	0.1462195
11	3	4	0.4012125	0.3410176	0.3245860
			0.1418767	0.1433937	0.1371549
12	3	5	0.3835971	0.3262388	0.3082678
			0.1297131	0.1323661	0.1242599
13	4	4	0.3879500	0.3305028	0.3132364
			0.1260315	0.1280590	0.1208380
14	4	5	0.3652429	0.3152385	0.2990205
			0.1280277	0.1281423	0.1215442
15	5	5	0.3548828	0.3050805	0.2908209
			0.1213273	0.1231452	0.1176504

Table 3.4: Mean and standard deviation for 500 samples, 50 points/sample, inspection length=1000

Distribution Number	Beta para. α	β	500 samples, 50 points/sample $y_{(n)} - y_{(1)}$	L_2	L_∞
1	1	1	0.9606377	0.9869183	0.9461884
			0.02678697	0.05241601	0.03054228
2	1	2	0.8647368	0.8769632	0.8492211
			0.06640976	0.07241678	0.06735622
3	1	3	0.7575122	0.7656219	0.7446620
			0.08175854	0.08317159	0.08069857
4	1	4	0.6524396	0.6578176	0.6419507
			0.09661906	0.09784351	0.09640397
5	1	5	0.5754476	0.5802976	0.5658817
			0.09441874	0.09505307	0.09422631
6	2	2	0.8526653	0.8549791	0.8264509
			0.05491866	0.06071470	0.05604649
7	2	3	0.7902669	0.7914479	0.7668523
			0.06711923	0.06866440	0.06509654
8	2	4	0.7182296	0.7182900	0.6953536
			0.07423409	0.07568630	0.07350819
9	2	5	0.6537800	0.6514293	0.6326223
			0.07270493	0.07497354	0.07229929
10	3	3	0.7596573	0.7579078	0.7319796
			0.06635577	0.06865339	0.06646436
11	3	4	0.7142692	0.7114468	0.6871443
			0.07119628	0.07270031	0.06859631
12	3	5	0.6659668	0.6621650	0.6410727
			0.07302258	0.07445586	0.07095007
13	4	4	0.6853173	0.6798982	0.6579420
			0.07195569	0.07189938	0.06915820
14	4	5	0.6571373	0.6518089	0.6302268
			0.06971710	0.07018052	0.06583591
15	5	5	0.6291794	0.6244845	0.6035707
			0.06556892	0.06700868	0.06248217

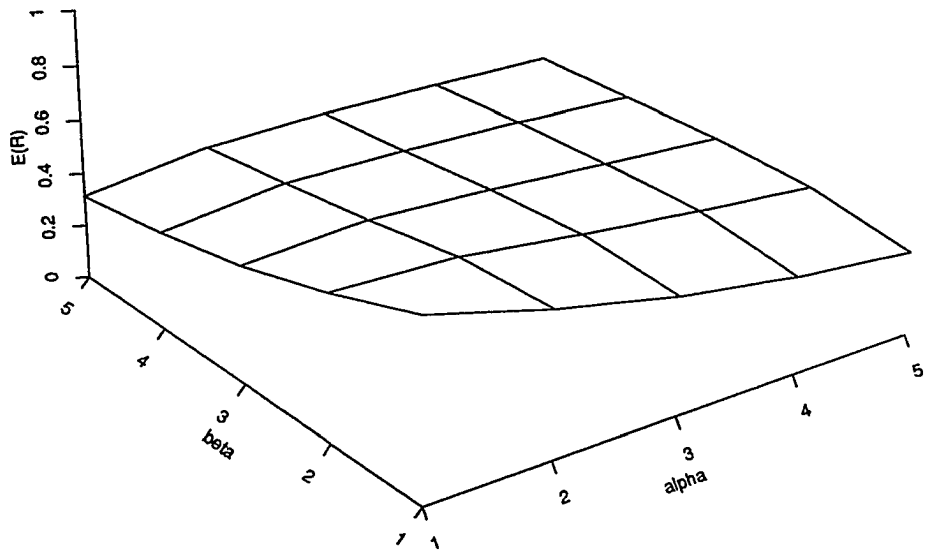


Figure 3.8: Mean of range from beta distribution for 5 measurement points

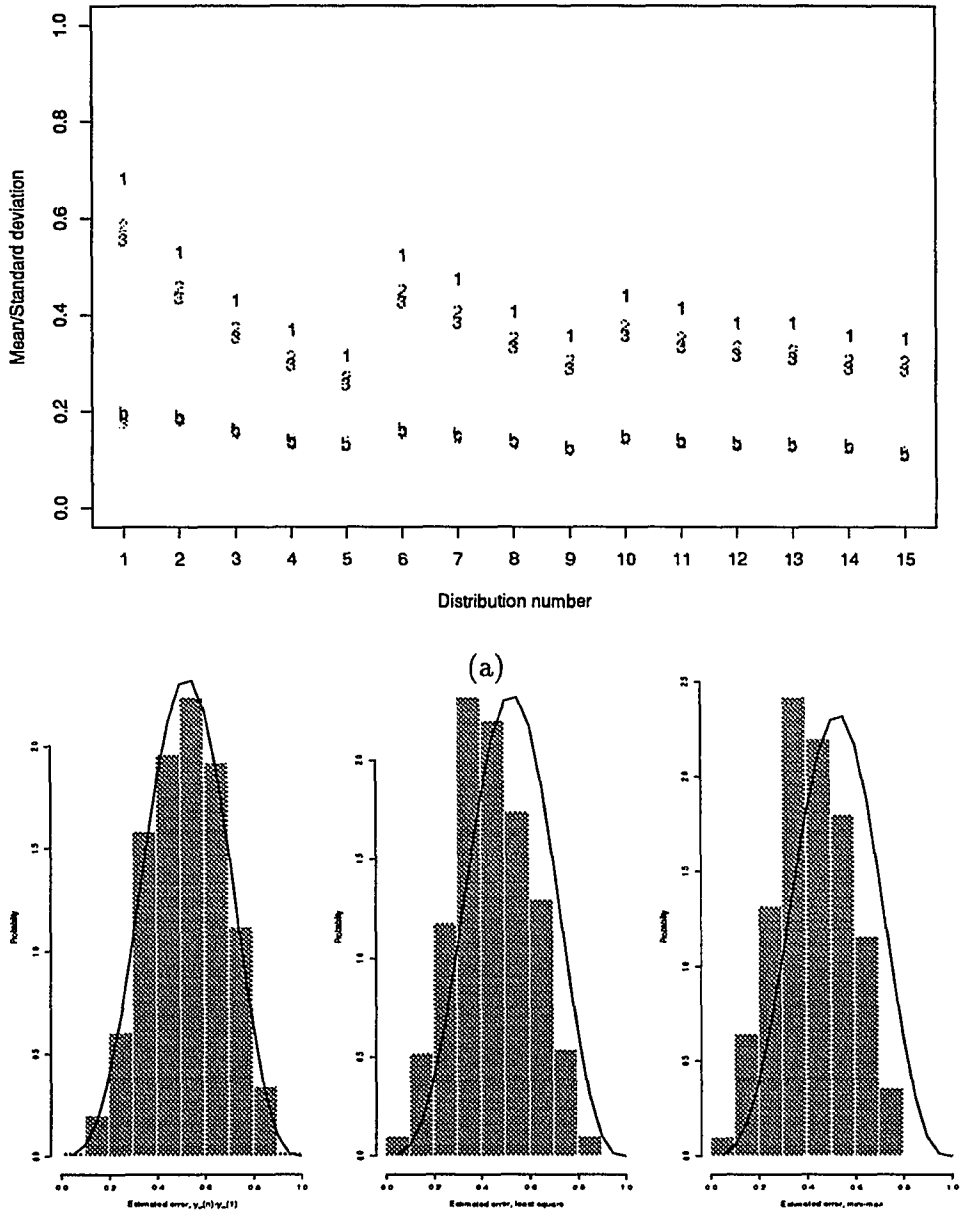
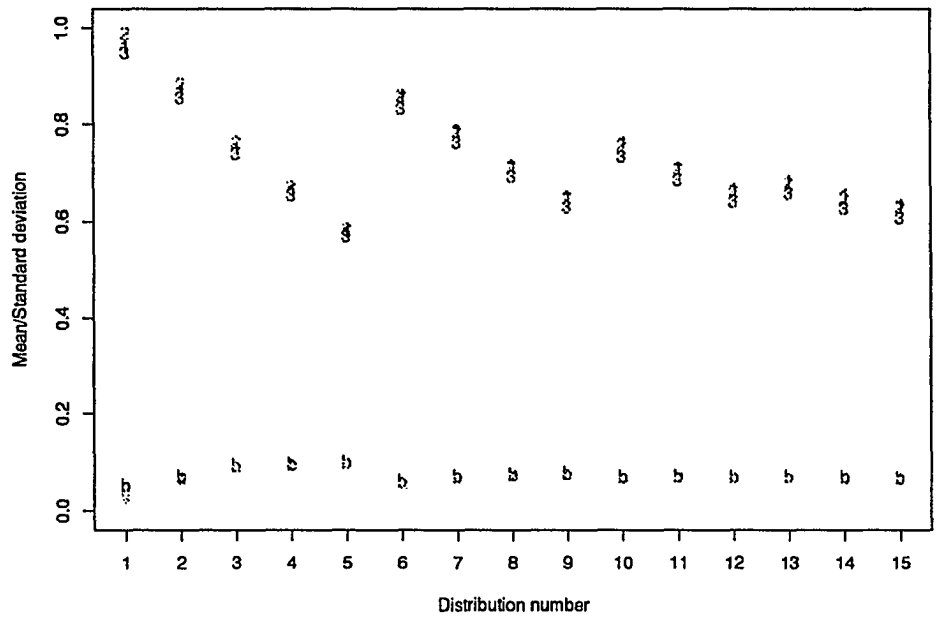
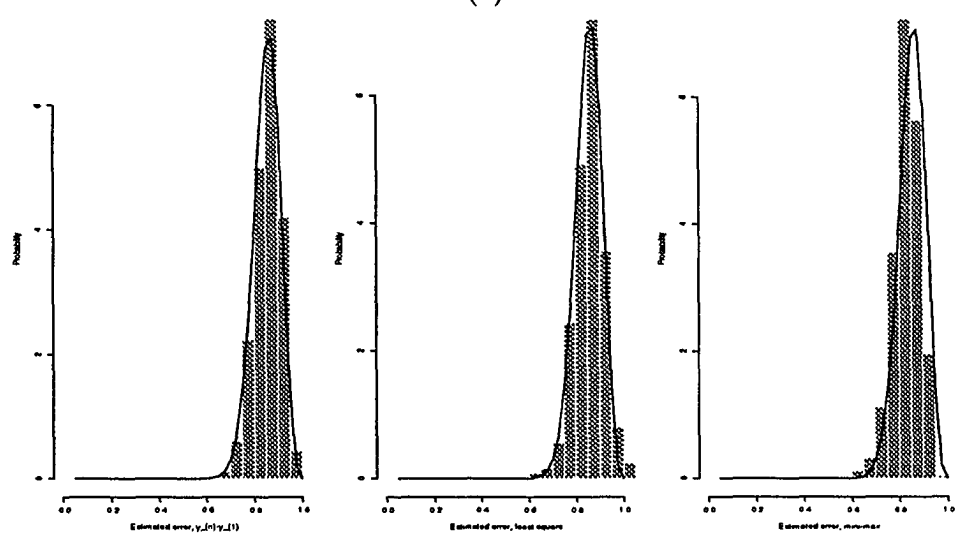


Figure 3.9: (a) Mean and standard deviation for 500 samples, 5 points/sample (1,2,3=mean, a,b,c=standard deviation of $y_{(n)} - y_{(1)}$, L_2, and L_∞ methods, respectively) (b) Estimated form errors histogram for beta(2,2) from 500 samples, 5 points/sample



(a)



(b)

Figure 3.10: (a) Mean and standard deviation for 500 samples, 50 points/sample (1,2,3=mean, a,b,c=standard deviation of $y_{(n)} - y_{(1)}$, L_2, and L_∞ methods, respectively) (b) Estimated form errors histogram for beta(2,2) from 500 samples, 50 points/sample

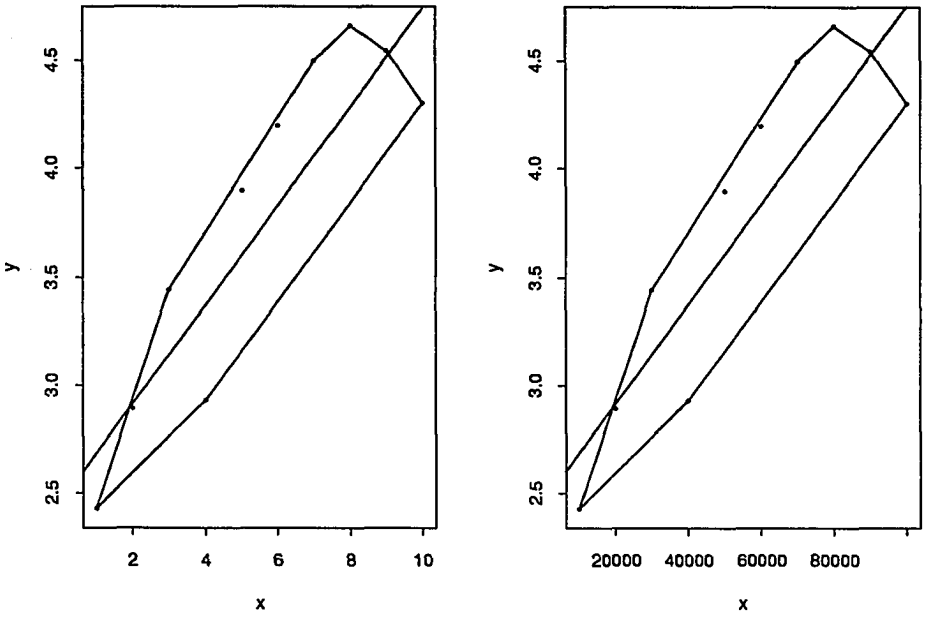


Figure 3.11: The L_∞ mean profile levels and convex hulls for the straightness data set 2 in Appendix B and x-axis scale 100000. The L_∞ mean profiles are $y = 2.456333 + 0.228667x$, $y = 2.456333 + 0.000023 * x$ and form errors are 0.857858, 0.88, respectively.

CHAPTER 4. SPATIAL STATISTICS FOR FORM ERROR ESTIMATION

In the previous discussion, we used the individual sample points to calculate the form error without applying any curve or surface fitting procedures to reconstruct the continuous profiles. Also, for the detectability graph, shown in Figure 3.5, we assumed that the measurement value at each point is independent of all other points. Thus, the question arises, is it necessary to consider the spatial dependence between sample points and apply surface fitting procedures to reconstruct the continuous profile from the discrete measurement information? More importantly, will the consideration of spatial dependence yield a better estimate of the true form errors?

Spatial prediction

Let the points on the surface we want to inspect be denoted by

$$\{Z(\mathbf{s}) : \mathbf{s} \in D\}$$

where $\mathbf{s}$ is a spatial location vector in $\mathbf{R}^2$ (two-dimensional reference datum plane). The index set D gives the extent of the region of the inspected surface. Data $\{Z(\mathbf{s}_1), ..., Z(\mathbf{s}_n)\}$ are measured at known locations $\mathbf{s}_1, ..., \mathbf{s}_n$.

Spatial dependence

In the previous chapter we characterized the surface by assuming that $Z(\mathbf{s}_1), ..., Z(\mathbf{s}_n)$ are independent of each other. However, the values of $Z(\mathbf{s})$ at locations in close proximity tend to be related to each other. The covariance function,

$$\text{cov}(Z(\mathbf{s}_i), Z(\mathbf{s}_j)) = C(\mathbf{s}_i, \mathbf{s}_j) \quad \text{for all } \mathbf{s}_i, \mathbf{s}_j \in D, \tag{4.1}$$

summarizes the relationship between the values $Z(\mathbf{s}_i)$ and $Z(\mathbf{s}_j)$ with locations $\mathbf{s}_i$ and $\mathbf{s}_j$. If $C(\mathbf{s}_i - \mathbf{s}_j)$ is a function only of $\| \mathbf{s}_i - \mathbf{s}_j \|$ (i.e., distance between the points), then $C(\cdot)$ is called isotropic. A random function $Z(\cdot)$ having covariance function (4.1) and

$$E(Z(\mathbf{s})) = \mu \quad \text{for all } \mathbf{s} \in D \tag{4.2}$$

is called second-order (or weak or wide-sense) stationary.

The variogram is an important parameter of geostatistics which also describes the relationship between values at two locations. The variogram

$$\text{var}(Z(\mathbf{s}_i) - Z(\mathbf{s}_j)) = 2\gamma(\mathbf{s}_i - \mathbf{s}_j) \quad \text{for all } \mathbf{s}_i, \mathbf{s}_j \in D \tag{4.3}$$

defines the variance of the difference of Z-values at separate locations. The function $\gamma(\cdot)$ is called a semivariogram. When the process is second-order stationary, the covariance function and the variogram are related by

$$\gamma(\mathbf{h}) = C(0) - C(\mathbf{h}).$$

Note that $C(0) = \text{var}(Z(\mathbf{s}))$ is the variance of $Z(\mathbf{s})$ at any location $\mathbf{s}$. Cressie (1993) points out that the variogram exists even for some processes that are not second-order stationary, and hence is more general than the covariance function. The majority

of machined surfaces are typically anisotropic (i.e., dependence between $Z(\mathbf{s})$ and $Z(\mathbf{s}+\mathbf{h})$ is a function of both the magnitude and the direction of $\mathbf{h}$ so that the variogram is no longer purely a function of the distance between 2 spatial locations) because they exhibit a pronounced lay or directional character. Sayles and Thomas (1977 and 1979) also point out the limitations of the covariance function and use the structure function

$$S(\mathbf{h}) = E\{[Z(\mathbf{s}) - Z(\mathbf{s}+\mathbf{h})]^2\}$$

to model the spatial dependence in their surface roughness studies.

Assuming that a process is stationary with respect to the mean (i.e., (4.2) holds), then $\mathrm{Var}(Z(\mathbf{s}+\mathbf{h}) - Z(\mathbf{s})) = E(Z(\mathbf{s}+\mathbf{h}) - Z(\mathbf{s}))^2$ and the variogram is estimated by

$$2\hat{\gamma}(d) = \frac{1}{n_d} \sum_{(i,j):\|\mathbf{s}_i - \mathbf{s}_j\| = d} (Z(\mathbf{s}_i) - Z(\mathbf{s}_j))^2, \tag{4.4}$$

where the summation is over all distinct pairs of locations in the sample that are separated by distance d, and n_d is the total number of pairs separated by d. Other robust estimations of variogram are also described in Cressie (1993). For irregularly spaced data, pairs of data with approximately the same distance apart may be grouped together. Finally, a smooth curve (e.g., linear, exponential, spherical, rational quadratic model) is fitted to the variogram estimates to obtain values of the variogram for all distances.

Kriging

Kriging is a stochastic processes prediction theory used to produce contour maps of surfaces derived from regularly or irregularly scattered points in a space. This theory incorporates the covariance function (or variogram) in the prediction process.

If we assume a process is intrinsic-stationary (i.e., (4.2) holds) and

$$\text{Var}(Z(\mathbf{s}_1) - Z(\mathbf{s}_2)) = 2\gamma(\mathbf{s}_1 - \mathbf{s}_2) \quad \text{for all } \mathbf{s}_1, \mathbf{s}_2 \in D,$$

then ordinary kriging can be used to predict the process. If the process is isotropic, then estimator (4.4) can be used. However, if the process is anisotropic, the variogram estimators are not only a function of distance but also a function of direction, i.e., $\gamma(\dot{\mathbf{n}}, \theta)$. Cressie (1993) gives an important note about the variogram estimators. If μ in (4.2) is not constant but in fact depends on the location $\mathbf{s}$, then the variogram estimators are estimating

$$2\gamma(\mathbf{h}) + (E(Z(\mathbf{s}+\mathbf{h})) - E(Z(\mathbf{s})))^2.$$

Thus, if the process is nonstationary, i.e., (4.2) does not hold, we can decompose the process into two structures

$$Z(\mathbf{s}) = \mu(\mathbf{s}) + \varepsilon(\mathbf{s})$$

where $E(Z(\mathbf{s})) = \mu(\mathbf{s})$ and $\varepsilon(\mathbf{s})$ is a zero-mean intrinsically stationary stochastic process with $\text{var}(\varepsilon(\mathbf{s}+\mathbf{h}) - \varepsilon(\mathbf{s})) = \text{var}(Z(\mathbf{s}+\mathbf{h}) - Z(\mathbf{s})) = 2\gamma(\mathbf{h})$. The large-scale variation $\mu(\mathbf{s})$ and small-scale variation $\varepsilon(\mathbf{s})$ are modeled as deterministic and stochastic processes. Examples of trend surfaces, $\mu(\mathbf{s})$, used in kriging are

$$\mu(\mathbf{s}) = a + c(x) + r(y), \quad \mathbf{s} = (x, y)' \tag{4.5}$$

where a is the overall trend, c is a column effect, and r is a row effect

$$\mu(\mathbf{s}) = \sum_{u+v \le p} a_{uv} x^u y^v, \quad \mathbf{s} = (x, y)', \tag{4.6}$$

where integer p is the order of the trend surface.

Expression (4.5) is the basis of median-polish kriging and (4.6) is the basis of universal kriging (Cressie, 1993). After the trend surface, $\mu(\mathbf{s})$, is estimated and removed, the residuals, $\varepsilon(\mathbf{s})$, are treated as stationary and a variogram can be fitted. Finally, the estimated residuals are combined with the trend surface to obtain estimates of the actual surface.

Venables and Ripley (1994) provide a software package in Splus (Becker, 1988) to implement universal kriging. The minimum mean square error unbiased predictor $\hat{Z}(x)$ is given by Ripley (1981) as

$$\hat{Z}(\mathbf{s}_0) = f(\mathbf{s}_0)^T \hat{\beta} + y^T k(\mathbf{s}_0). \tag{4.7}$$

The computation procedure is summarized as follows.

1. Form $K = [C(\mathbf{s}_i, \mathbf{s}_j)]$
2. Find L such that $LL^T = K$ where LL^T is the Cholesky decomposition of K and L is lower triangular.
3. Form $F = \begin{bmatrix} f_1(\mathbf{s}_1) & \cdots & f_P(\mathbf{s}_1) \\ \vdots & & \\ f_1(\mathbf{s}_1) & \cdots & f_P(\mathbf{s}_N) \end{bmatrix}$ and $Z_N = \begin{bmatrix} Z(\mathbf{s}_1) \\ \vdots \\ Z(\mathbf{s}_N) \end{bmatrix}$
   where f is polynomial function for the trend surface and P is the number of coefficients $P = (p+1)(p+2)/2$ where p is the order of the trend surface.
4. Solve $L^{-1} Z_N = L^{-1} F \hat{\beta}$ by least squares for $\hat{\beta}$.
5. Form $W_N = [Z(\mathbf{s}_i) - f(\mathbf{s}_i)^T \hat{\beta}]$
6. Find y such that $L(L^T y) = W_N$

7. Predict $Z(\mathrm{s}_0)$ by $f(\mathrm{s}_0)^T\hat{\beta} + y^T k(\mathrm{s}_0)$, where $k(\mathrm{s}_0) = [C(\mathrm{s}_0, \mathrm{s}_i)]$ and

$$var[Z(\mathrm{s}_0) - \hat{Z}(\mathrm{s}_0)] = C(\mathrm{s}_0, \mathrm{s}_0) - \| e \|^2 + \| g \|^2$$

where $Le = k(\mathrm{s}_0)$, $R^T g = f(\mathrm{s}_0) - (L^{-1}F)^T e$, and R is the orthogonal reduction of $L^{-1}F$.

We will not consider the extrapolation of the fitted surface which is outside the convex hull, CH, formed by the sample points in $X-Y$ plane, i.e., $\hat{Z}(\mathrm{s})$ with $\mathrm{s} \in CH$ only. The detailed procedure to calculate the flatness error by using universal kriging is described as follows.

1. Take random samples $Z(\mathrm{s}_1), ..., Z(\mathrm{s}_N)$ from the inspected surface.

2. Use Venables and Ripley's programs to calculate $\hat{Z}(\mathrm{s})$ and σ_E^2 within D. Note that the grid size specified in their program in X and Y directions to obtain the surface coordinates will affect the approximation of the flatness errors of the predicted surface. That means the flatness error calculated for a predicted surface is less or equal to its true flatness error.

3. Obtain the 2D convex hull only considering X and Y coordinates of the sample points. QHULL (Barber and Huhdanpaa, 1994) provides a program to find 2D and 3D convex hull for the given points.

4. Remove the fitted surface outside the convex hull of $X - Y$ plane.

5. Calculate the flatness error of the predicted surface (X-Y coordinates) inside the convex hull, CH.

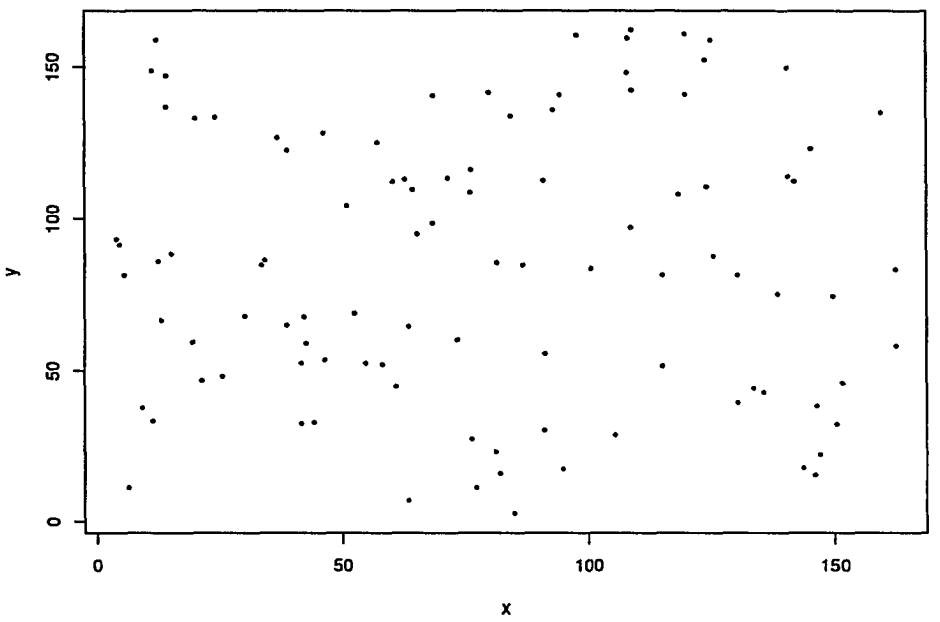


Figure 4.1: Sample points (100) taken from the bored surface (bocl1)

Random points from a bored surface

This section illustrates the use of universal kriging to calculate the flatness error for the manufactured surface (bocl1) from a boring process (Stout et al., 1990). The region D for this bored surface is $(x, y) \in ([0, 163], [0, 163])$ which is shown in Figure C.1. The surface data are taken from a uniform rectangular grid giving a total of 26,896 points (164×164). The true flatness error is 41.164028 for these 26,896 points which is unknown to the inspector. First we take 100 sample points from this surface by using a random sampling method. The scatter plot of these points in X-Y plane is shown in Figure 4.1. If we only consider the individual points, the flatness form error as calculated in the previous chapter is 29.306071. Next we use universal kriging to find a predicted surface for these points.

The variogram must satisfy a property called conditional negative-definiteness, i.e.,

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j 2\gamma(\mathbf{s}_i - \mathbf{s}_j) \leq 0,$$

for any finite number of spatial locations $\{\mathbf{s}_i : i = 1, ..., m\}$ and real numbers $\{a_i : i = 1, ..., m\}$ satisfying $\sum_{i=1}^{m} a_i = 0$. The variogram estimators, e.g., $\hat{\gamma}(d)$, cannot be used for kriging because they are not necessarily conditionally negative-definite which can result in embarrassing negative mean-squared errors of prediction. It is suggested that a variogram model be selected from among a parametric family of variograms which best fits the data (Cressie, 1993). Figure 4 shows correlation plots for the data and the residuals from first and second order surfaces, together with covariance functions fitted by eye (the fitting of correlation (variogram) function is a subjective process). The covariance functions[1] used are the exponential model

$$C(h, a_e) = e^{-\|\mathbf{h}\|/a_e} \tag{4.10}$$

and wave model

$$C(h, a_w) = a_w \frac{\sin(\| \mathbf{h} \| /a_w)}{\| \mathbf{h} \|}. \tag{4.11}$$

The parameters, a_e and a_w, must be estimated for each model. If $\gamma(\mathbf{h}) \to c_0 > \mathbf{0}$ as $\mathbf{h} \to \mathbf{0}$, then c_0 is called the nugget effect which is caused by measurement error. Since

[1]The corresponding variogram models are the exponential model

$$\gamma(h, a_e, b_e) = b_e \left[1 - e^{-\|\mathbf{h}\|/a_e}\right] \tag{4.8}$$

and wave model

$$\gamma(h, a_w, b_w) = b_w \left[1 - a_w \frac{\sin(\| \mathbf{h} \| /a_w)}{\| \mathbf{h} \|}\right]. \tag{4.9}$$

The parameters, (a_e, b_e) and (w_a, w_b) must be estimated for each model.

we consider the measurement error to be negligible, the nugget effect is not included in these covariance functions. That is, the correlation is 1.0 when the distance is 0.0 which will result in the krigged surfaces going through those measurement points. Introducing a trend surface makes little difference to these correlograms. Figure 4.3 and 4.4 show the predicted surfaces and the prediction standard errors for these two models fitted in Figure 4. As can be seen from the two surfaces, the choice between different covariance models (exponential and wave models) makes a big difference in prediction.

Although the predicted surfaces are continuous, we use a 55×55 sample grid within the region D to represent this predicted surface. Since we do not consider the extrapolation of the predicted surface, we find the convex hull for the sample points in region D which is shown in Figure 4.5. Only the grids inside the convex hull are calculated by the program to determine the flatness error. The flatness error is 25.544803 for Figure 4.3 (exponential model) if we only calculate the predicted surface within the convex hull region, which is smaller than 29.306071 obtained by using individual points only. Because we use a rectangular grid to represent the predicted surface and calculate the flatness error, the results are biased low because we may miss the minimum and maximum points. Figure 4.4 gives a large flatness error of 72.275202 which is much larger than the true error 41.164028 calculated from all 26,896 points.

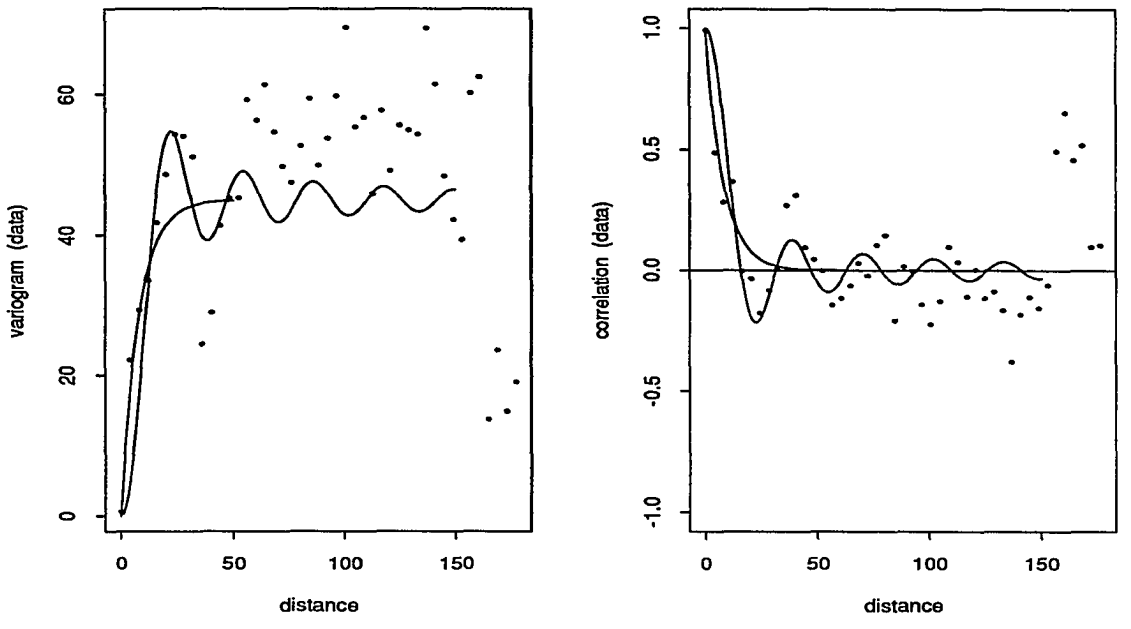
Discussion

Detrending the data is an important issue in kriging. Universal kriging is limited to polynomial trend surfaces. Cressie (1993) presents median-polish kriging which

provides a more flexible and statistically resistant method of spatial prediction than universal kriging. Another more serious problem is the choice of a variogram model (covariance function) which can make a big difference in prediction. We see from the previous example that incorrectly fitting the wave model results in an overshoot for predicted surfaces. Also, the variogram fitting should use only up to half the maximum possible lag and then only using lags for which $n_d > 30$ (Cressie, 1993). Thus, the empirical variogram fitted from the sample data usually needs a large number of samples. It is also important to have a good fit for the variogram at small distances between data points. In sampled data analysis, nothing can be said about the variogram at lag distances smaller than $\min\{\| \mathbf{s}_i - \mathbf{s}_j \| : 1 \leq i < j \leq N\}$. For a small number of sample points, this poses a significant problem.

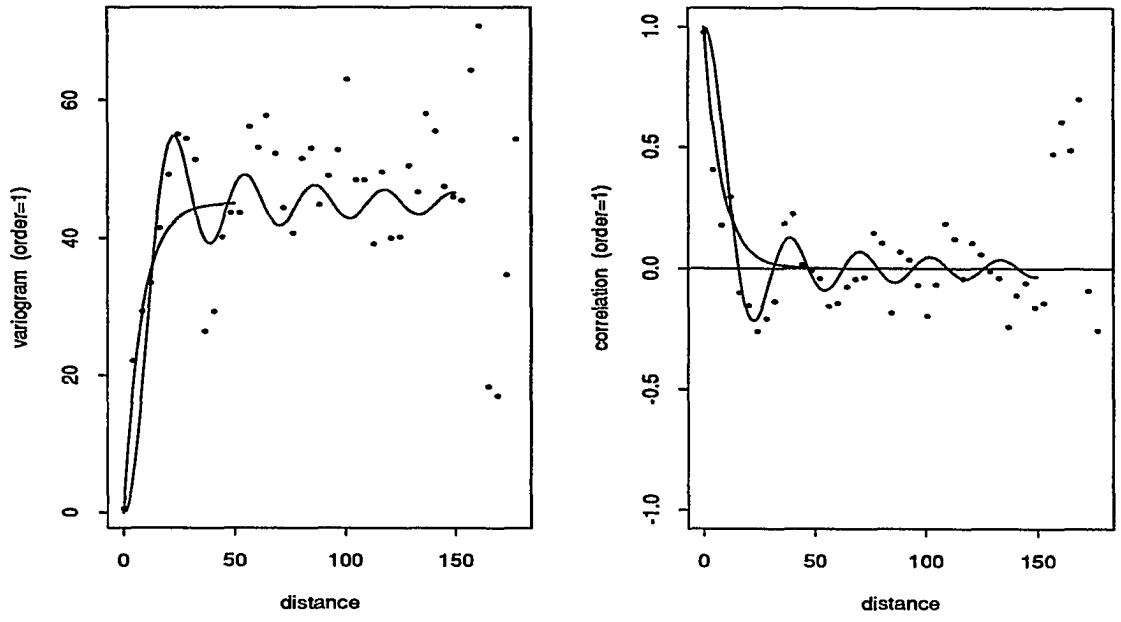
After performing universal kriging on a number of surfaces, we found the following phenomenon. The variogram determines whether the predicted surface falls inside or outside the 3-D convex hull of the data points. If the slope of the variogram approaches zero as the distance approaches zero, kriging will return values which may be outside the 3-D convex hull. If the variogram has a slope which is sufficiently greater than zero when the distance is zero, the resulting interpolated value will lie within the 3-D convex hull. Therefore, the commonly adapted exponential model (the spatial dependence getting smaller as the distance increases) will always result in a predicted surface within the 3-D convex hull. This is an unfavorable situation since this provides no additional information to standard interpolation. Based on the large number of points required and difficulty in identifying the correct empirical variogram model (covariance function) from small number of samples, there appears to be no significant advantage in applying the kriging approach to determine the form

error.

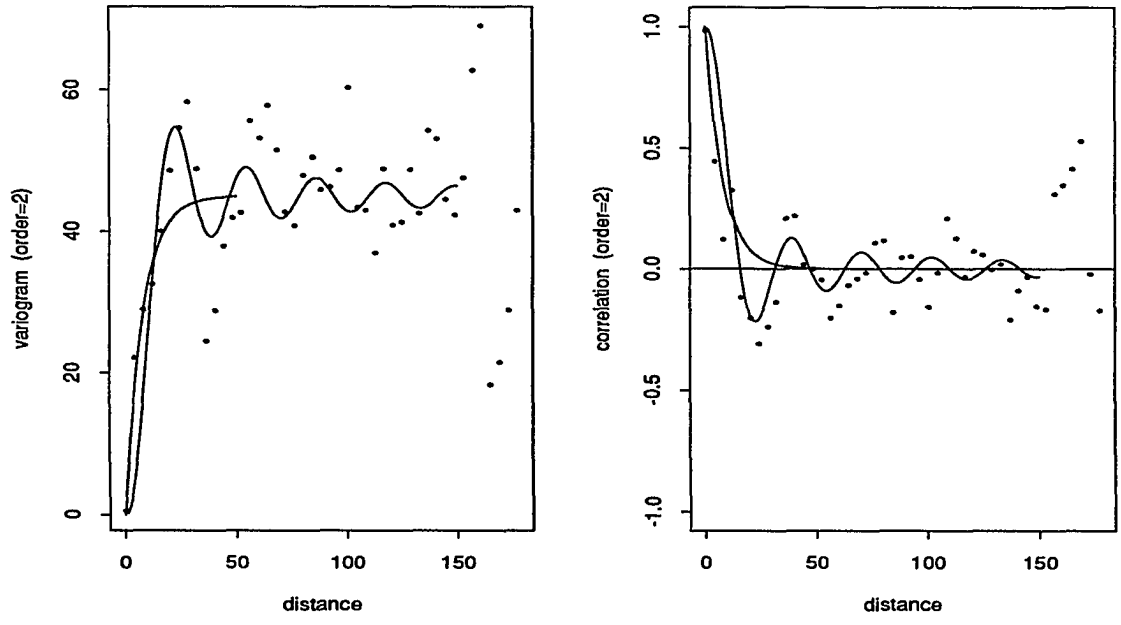


(a)

Figure 4.2: The variograms and correlograms. (a) Data. (b) Residuals from first order surface. (c) Residuals from quadratic surface. The fitted covariances are exponential model with $a_e = 8$ and wave model with $a_w = 5$.

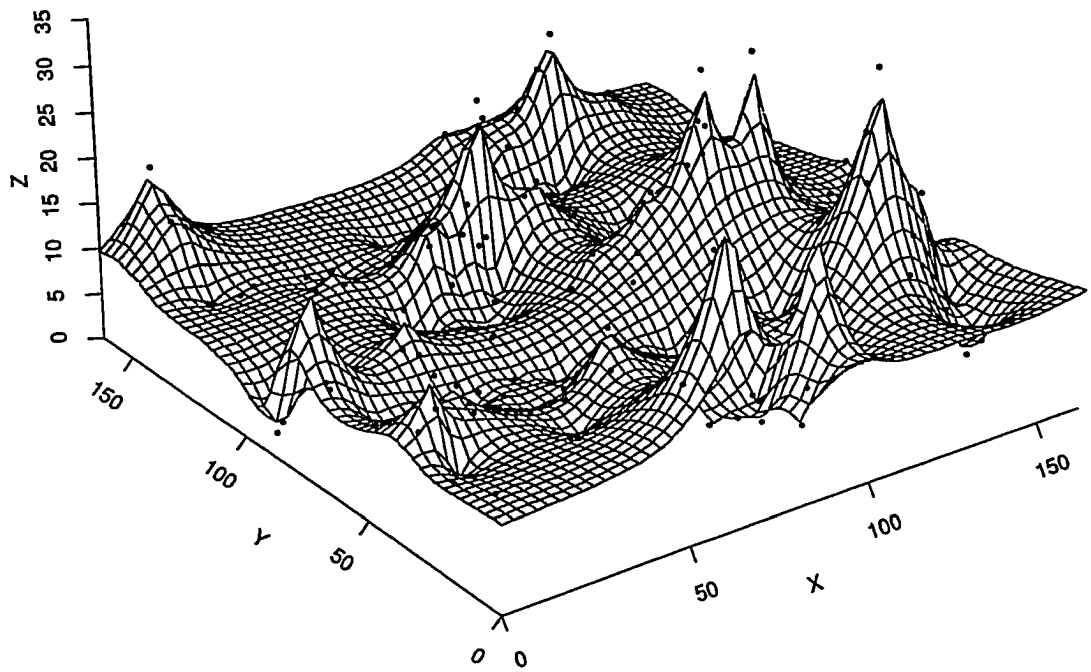


(b)



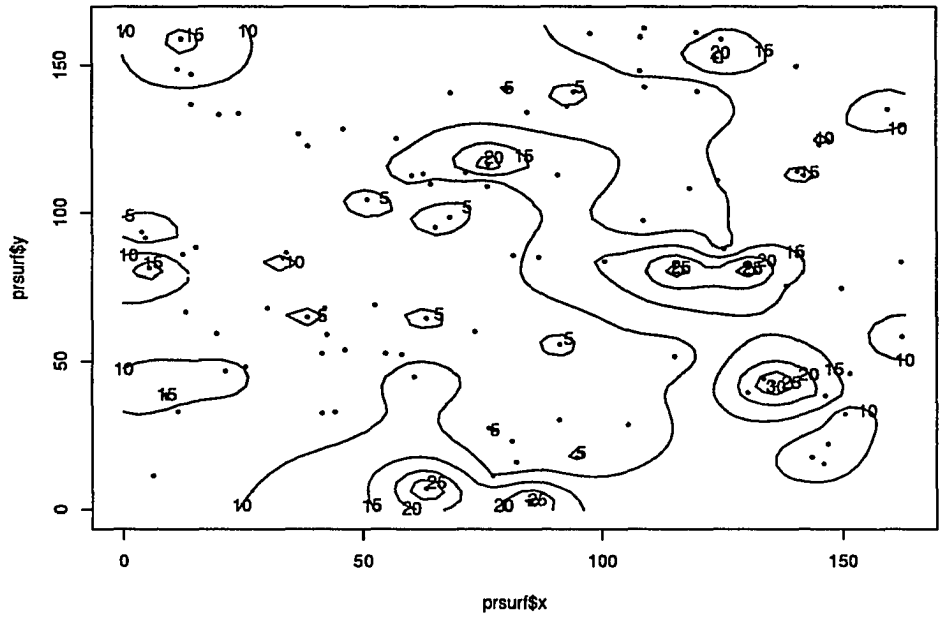
(c)

Figure 4.2: (Continued)

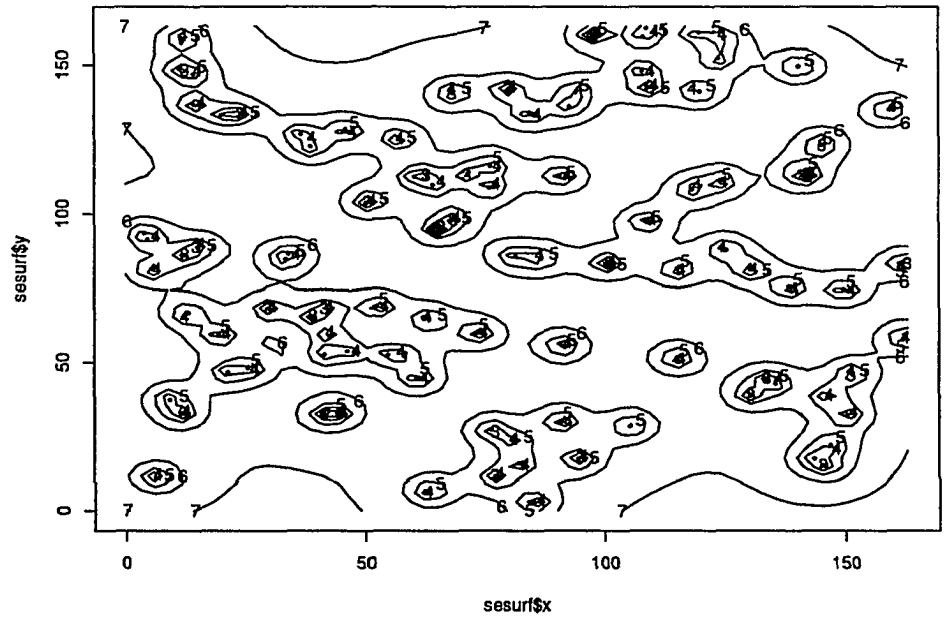


(a)

Figure 4.3: Predicted surface (a), (b) and standard error of the prediction error (c) for exponential model with $a_e = 8$.

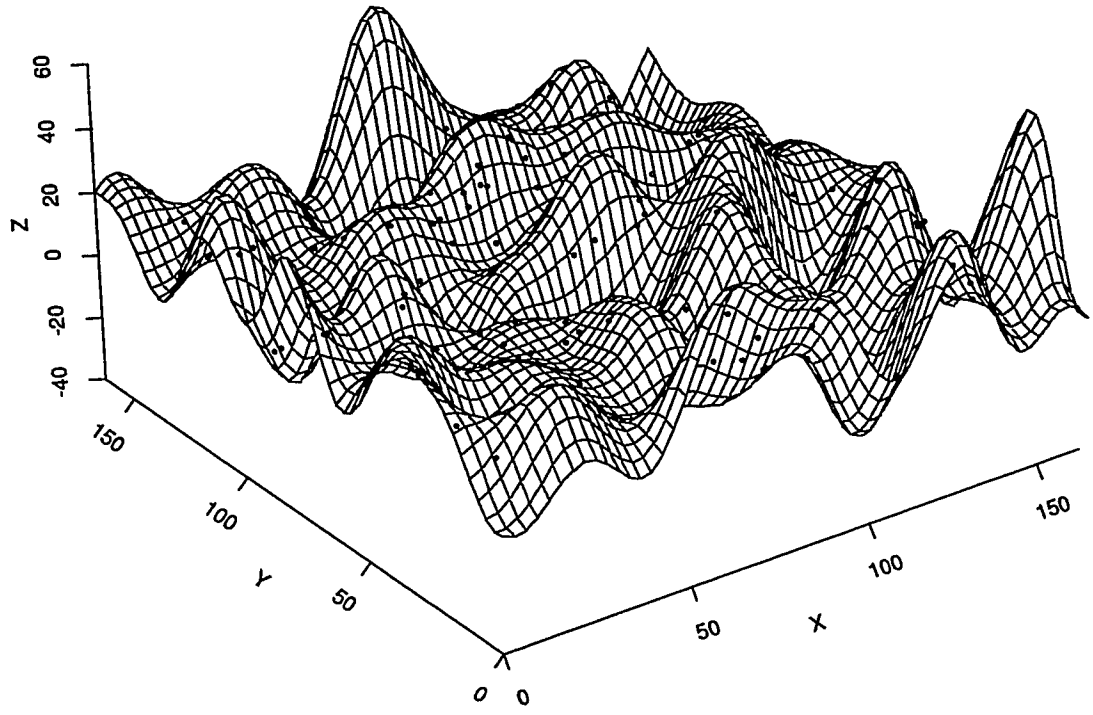


(b)



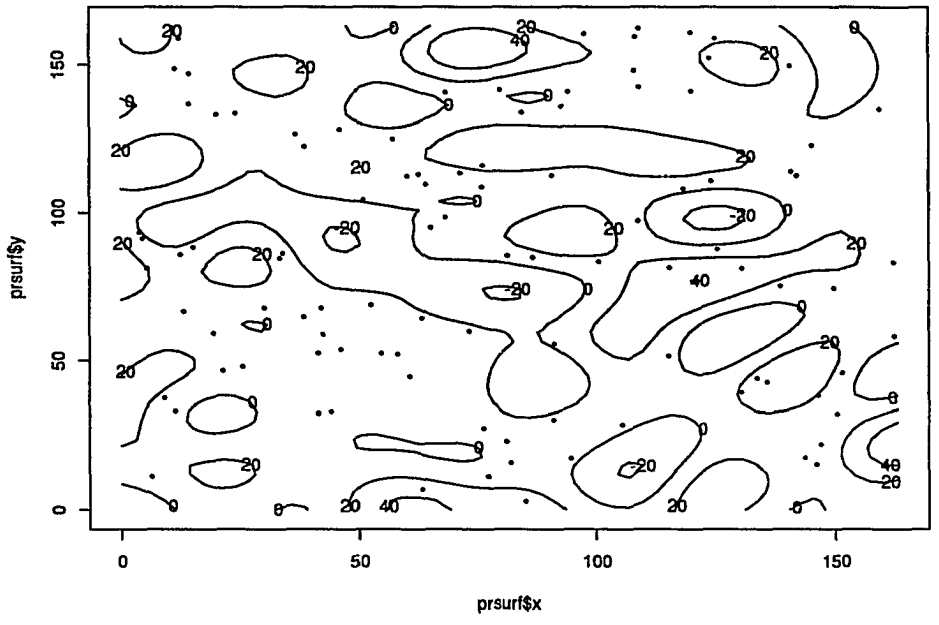
(c)

Figure 4.3: (Continued)

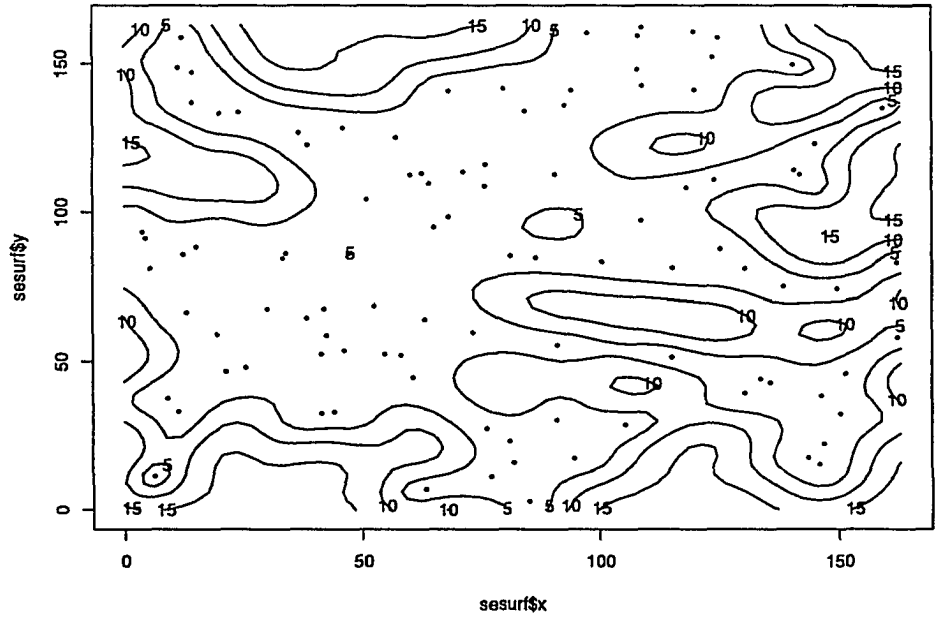


(a)

Figure 4.4: Predicted surface (a), (b) and standard error of the prediction error (c) for wave model with $a_w = 5$.



(b)



(c)

Figure 4.4: (Continued)

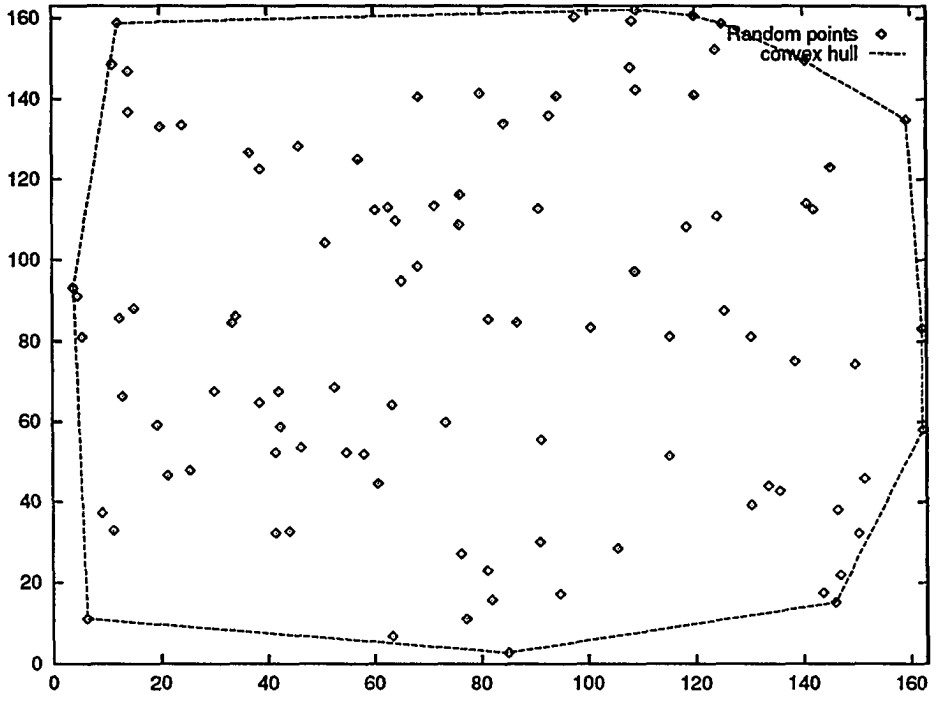


Figure 4.5: The convex hull, CH

CHAPTER 5. UNIFORM SAMPLING AND OPTIMAL INTERPOLATION

Uniform sampling

The only sampling method discussed to this point is uniform random sampling. Ripley (1981) evaluated the performance of uniform random sampling, stratified random sampling, and uniform (systematic) sampling on the estimation of the mean value within region D. He calculated variance of the error of mean value estimation, $N\text{var}[\sum Z(\mathbf{s}_i)/N - \int_D Z(\mathbf{s})d\mathbf{s}/\text{area of } D]$, and found that if low frequencies are dominant (corresponding to strong local positive correlation), both stratified random and systematic sampling should do well relative to uniform random sampling. He further concluded that uniform (systematic) sampling should be the best with smaller error variance unless the process has strong periodicity with a wavelength corresponding to the basic sampling interval along either axis or with wavelength along a diagonal. In this chapter we confine our discussion to uniform sampling strategies and present an optimal interpolation method for surface reconstruction from a small number of discrete sample points. The performance of uniform sampling and optimal interpolation, along with universal kriging, for flatness error estimation will be investigated in the next chapter.

Interpolation methods

Many interpolation methods exist in the literature. Watson (1992) extensively reviewed the existing interpolation methods. He classifies these methods into five categories:

1. Distance-based methods. These methods assume that each datum has local influence that diminishes with distance and becomes negligible beyond a limiting radius.

2. Fitted function methods. Some examples include Lagrange interpolation, collocation, minimum curvature splines, kriging, and relaxation surfaces.

3. Triangle-based methods. For example, Akima's (1978) method uses a fifth-degree polynomial interpolation function in x and y defined in each triangular cell.

4. Rectangle-based methods. These methods use rectangular grids for sample data, for example, bilinear, Hermite, Bezier, B-spline, and tension patches, Taylor interpolation, and Fourier surfaces.

5. Neighborhood-based methods. These methods are closely related to distance-weighted methods.

Among these methods, kriging is synonymous with "optimum prediction" or "optimally predicting" (p. 119, Cressie, 1993) in the stochastic prediction content, which we already discussed in the previous chapter. Another method, spline interpolation, is also considered to be an optimal theoretical solution to the estimation

of functions from limited data (Powell, 1981) in approximation theory or numerical methods. In the following presentation, we show that the Shannon sampling functions are the optimal interpolation for surface reconstruction from a limited set of sample points.

Shannon sampling theory

An important family of mathematical techniques used in communication engineering and information theory are based on Shannon sampling theory. We can treat the surface profiles as signals in a time domain where distance represents time. A finite-energy signal, $z(\mathrm{s})$, is said to be band-limited if its amplitude spectrum (its Fourier transform) vanishes outside an interval of the form $(-W, W)$, where W is called the bandwidth of the signal and $\mathrm{s} \in \mathbf{R^1}$ (one-dimensional reference datum line), i.e.,

$$z(\mathrm{s}) = \frac{1}{\sqrt{2\pi}} \int_{-W}^{W} Z(\omega) e^{i\mathrm{s}\omega} d\omega,$$

where $i = \sqrt{-1}$ and Z satisfies $\int_{-W}^{W} |Z(\omega)|^2 \, d\omega < \infty$. Since $Z(\omega)$ is zero for $|\omega| > W$, we can replicate it to form a periodic function in the frequency domain with period $2W$. This periodic function can be expressed as a Fourier series (Marks, 1991), i.e.,

$$Z(\omega) = \begin{cases} \sum_{N=-\infty}^{\infty} c_N e^{-i\pi N\omega/W}, & |\omega| < W \\ 0, & |\omega| \geq W, \end{cases} \tag{5.1}$$

where the Fourier coefficients are

$$c_N = \frac{1}{2W} \int_{-W}^{W} Z(\omega) e^{i\pi N\omega/W} d\omega = \frac{1}{2W} z\left(\frac{N}{2W}\right)$$

Substituting c_N into (5.1) and performing an inverse transform gives the sampling theorem series,

$$z(\mathrm{s}) = \frac{1}{\pi} \sum_{N=-\infty}^{\infty} z\left(\frac{N}{2W}\right) \frac{\sin[\pi(2W\mathrm{s} - N)]}{2W\mathrm{s} - N}. \tag{5.2}$$

If the surface signals (profiles) are band-limited functions, this sampling theorem states that it is possible to recover the intervening values with full accuracy (Marks, 1991). In other words, the sample set can be fully equivalent to the complete set of signal values. The minimum sampling rate (i.e., number of samples per unit distance) is equal to two samples per period of the highest frequency component of the signal. If W is the signal's bandwidth (the highest frequency component), then the signal $z(\mathrm{s})$ can be reconstructed from the samples by (5.2).

Equation (5.2) is called the cardinal series of Shannon sampling. This theorem was extended to n variables as follows (Zayed, 1993). Let $z(\mathrm{s})$ be a function of n real variables, where $\mathrm{s} \in \mathbf{R^n} = \{x_1, x_2, ..., x_n\}$, whose n-dimensional Fourier integral exists and is identically zero outside an n-dimensional rectangle and is symmetrical about the origin; i.e., $Z(\omega_1, \omega_2, ..., \omega_n) = 0, \quad |\omega_k| > |W_k|, \quad k = 1, 2, ..., n$. Then, the n dimensional signal can be recovered from the samples by

$$z(\mathrm{s}) = \sum_{m_1=-\infty}^{\infty} \cdots \sum_{m_n=-\infty}^{\infty} z\left(\frac{\pi m_1}{W_1}, ..., \frac{\pi m_n}{W_n}\right)$$

$$\frac{\sin(W_1 x_1 - m_1\pi)}{W_1 x_1 - m_1\pi} \cdots \frac{\sin(W_n x_n - m_n\pi)}{W_n x_n - m_n\pi} \tag{5.3}$$

where m_k is the sample number for dimension k. Marks (1991) and Zayed (1993) give detailed historical background and extensions to this theorem.

B-spline interpolation

The spline function is known as an interpolation function which is useful in constructing a smooth function from a given data sequence. B-spline interpolations are piecewise polynomials and have several advantages over other polynomial fits. As mentioned before, they also provide optimal solutions to the estimation of functions from limited data. However, the use of B-spline representations has had limited application in the signal processing field. The main reason for this lack of acceptance is because the conventional approach to B-spline interpolation or approximation is computationally expensive involving explicit matrix inversions and multiplications (Unser, 1993). This is also the reason that the simpler algorithm, cubic spline interpolation, instead of higher order splines is implemented in most computer software packages. The following derivation gives the B-spline interpolation as a linear combination of the sample points. Therefore, B-spline interpolation can be performed in real-time, synchronized to the successively given sampling points as required in a signal processing environment. From this derivation, we can relate B-spline interpolation of a general degree to Shannon sampling theory.

Let the sample point sequence on the space axis be $\{\mathrm{s}_n\}_{n=-\infty}^{\infty}(\mathrm{s}_n = nh, n = 0, \pm 1, \pm 2, ...)$, then the B-spline interpolation function determined by the sample points $z(\mathrm{s}_n)$, $n = 0, \pm 1, \pm 2, ...$ can be represented as (Kamada, et al., 1991)

$$z(\mathrm{s}) = \sum_{n=-\infty}^{\infty} z(\mathrm{s}_n)\ \phi[s]_0^m(\mathrm{s} - nh). \tag{5.4}$$

The spline sampling function $\phi[s]_n^m$ is invariant to the translation by h, i.e.,

$$\phi[s]_n^m(s) = \phi[s]_0^m(s - nh),\ \ n = 0, \pm 1, \pm 2, ...,$$

and it is given by

$$\phi[s]_n^m(s) = \sum_{l=-\infty}^{\infty} \beta^m(l-n)\ \phi[b]_l^m(s), \quad n = 0, \pm 1, \pm 2, ... \tag{5.5}$$

where

$$\phi[b]_l^m(s) = \frac{1}{h^m} \sum_{p=0}^{m} \frac{(-1)^p (s - (l + \frac{m}{2} - p)h)_+^{m-1}}{p!(m-1)!}, \quad l = 0, \pm 1, \pm 2, ... \tag{5.6}$$

and

$$(s-a)_+^{m-1} = \begin{cases} (s-a)^{m-1}, & s > a \\ 0, & s <= a \end{cases}$$

(5.6) is $(m-2)$-times continuously differentiable,

$$\beta^m(p) = h \int_{-h/2}^{h/2} B_f^m(f) e^{i2\pi fph} df, \quad p = 0, \pm 1, \pm 2, ...,$$

and

$$B_f^m(f) = \frac{1}{\sum_{q=-\lfloor \frac{m-1}{2} \rfloor}^{\lfloor \frac{m-1}{2} \rfloor} \phi[b]_0^m(qh) e^{-i2\pi fph}},$$

$\lfloor x \rfloor$ is the maximum integer not exceeding x.

If we rewrite the Shannon sampling series (5.2) as an interpolation function with the sample points,$s_k = \frac{k\pi}{W}$, $k = 0, \pm 1, \pm 2, ...$, as its coefficients, then (5.2) becomes

$$z(s) = \sum_{k=-\infty}^{\infty} z(s_k) \frac{\sin[W(s - s_k)]}{W(s - s_k)} \tag{5.7}$$

The function

$$S_k(s) \equiv \frac{\sin[W(s - s_k)]}{W(s - s_k)} \tag{5.8}$$

is called a sampling function or ideal (sinc) interpolation function (Unser, 1993), which is shown in Figure 5.1 ($W = 2, s_k = 0$). Unser, et al. (1993) showed that

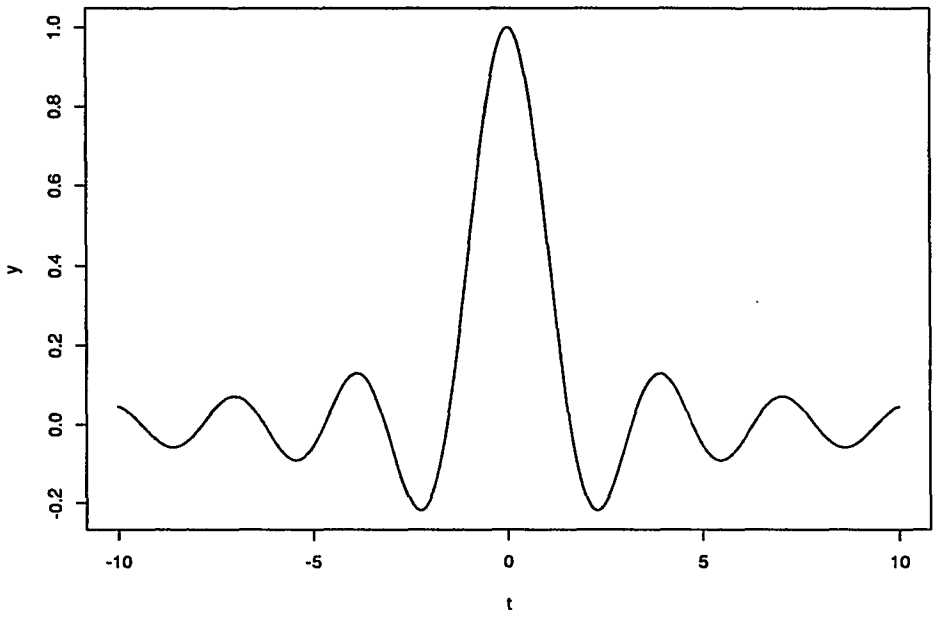


Figure 5.1: Ideal sinc interpolation function

as the order m of the spline sampling function (5.5) approaches infinity, $\phi[s]_n^m(s)$ converges to (5.8), the ideal sinc interpolation functions, i.e.,

$$\lim_{m\to\infty} \phi[s]_n^m(s) \longrightarrow S_k(s). \tag{5.9}$$

Thus, the Shannon sampling function is equivalent to an infinite order B-spline interpolation function. This infinite order B-spline interpolation function has an important meaning in the approximation of bandlimited functions which is discussed in the optimal interpolation section. Since we only use a small number of sample points to approximate the surface profile, the errors associated with this approximation are discussed as follows.

Undersampling

The cardinal series of Shannon sampling requires an infinite number of samples. If a signal has finite energy, it must asymptotically approach zero at $t = \pm\infty$. Therefore, there is always an interval of duration T outside of which the samples are negligibly small. If we sample over this interval at the Nyquist rate, $2W$, then a total of $S_T = 2WT$ samples are needed to characterize the signal (Marks, 1991). The choice of T is dependent by the truncation error one can tolerate. Thus, the knowledge of the frequency bound determines the minimum rate at which the signal needs to be sampled in order to reconstruct it completely. However, when we apply this theorem to the CMM context, we typically do not have knowledge *a priori* about the frequency information of a manufactured surface. Also, we typically take a relatively small number of sample points. This is described as an undersampling situation. Therefore, we commit two types of errors when we apply this theorem to the analysis of CMM measurement points, namely, truncation error and aliasing error (Marks, 1991). Truncation error occurs when only a limited number of samples are measured instead of the S_T number of samples needed. Aliasing error is the result of using a smaller bandlimit B ($B < W$), so that,

$$\epsilon_A = z(\mathrm{s}) - z_B(\mathrm{s}) = z(\mathrm{s}) - \frac{1}{\pi} \sum_{N=-\infty}^{\infty} z\left(\frac{N}{2B}\right) \frac{\sin[\pi(2B\mathrm{s} - N)]}{2B\mathrm{s} - N},$$

where $z_B(\mathrm{s})$ is the interpolated profile by using smaller bandlimit B.

If we look at the surface plots in Appendix C, we notice that in these surfaces the high frequency component is associated with low amplitude and the low frequency component is associated with high amplitude. This phenomenon has been observed in other experimental evidence (Sayles and Thomas, 1978). These low frequency

components with high amplitudes play an important role in the determination of form errors. Thus, the high frequency component is not as significant in terms of form error estimation. Therefore, the aliasing error is not as severe as the truncation error. The following theorem provides optimal theoretical solutions with a minimized error bound for the estimation of functions from small number of sample points.

Optimal interpolation

Powell (1981) formulated the optimal interpolation problem as follows. The sample points $z(\mathrm{s}_n)$, $n = 1, 2, ..., m$ of a function z in $C^{(k+1)}[a, b]$ are given, where $C^{(k+1)}[a, b]$ is a set of (k+1) differentiable continuous functions that are defined on the interval $[a, b]$ of the real line. If it is known that $\max_{a \le s \le b} |z^{(k+1)}(s)|$ is not very large, and if we want to estimate $z(\xi)$, where ξ is any point of $[a, b]$, then one may make an approximation of the form

$$z(\xi) \approx \sum_{i=1}^{m} w_i z(s_i)$$

where the multipliers $\{w_i : i = 1, 2, ..., m\}$ are such that the approximation is exact when z is in a space of all real polynomials of degree at most k. In this case the Peano kernel theorem shows that there is a real number c, that is independent of z, such that the bound

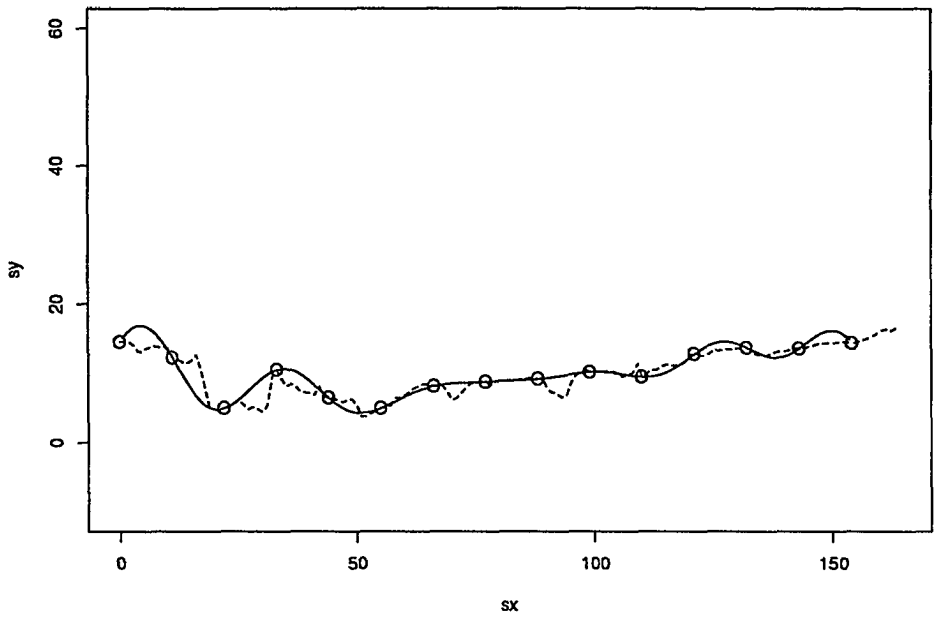
$$|z(\xi) - \sum_{i=1}^{m} w_i z(\mathrm{s}_i)| \le c \quad \max_{a \le s \le b} |z^{(k+1)}(\mathrm{s})|, \quad z \in C^{(k+1)}[a, b], \qquad (5.10)$$

is satisfied. Thus, the optimal interpolation problem is to find a minimum value for c. Let $s(\xi)$ be the optimal estimate

$$s(\xi) = \sum_{i=1}^{m} w_i(\xi) z(\mathrm{s}_i), \quad a \le \xi \le b, \qquad (5.11)$$

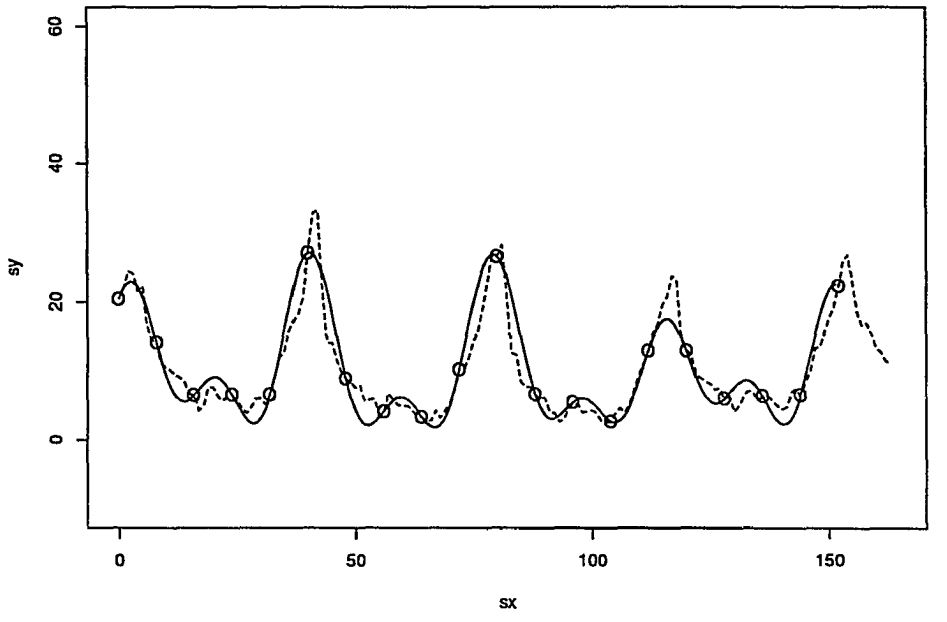
of $z(\xi)$. Powell proved that the optimal multipliers $\{w_i(\xi); i = 1, 2, ..., m\}$ are unique for each ξ and the approximation is a B-spline of degree k that has $(m-k-1)$ knots whose positions are independent of z. Because the optimal interpolation procedure can be applied for all values of ξ in $[a, b]$, the function (5.11) can be regarded as an approximation to the function $\{z(\mathbf{s}) : a \leq \mathbf{s} \leq b\}$.

Since the bandlimited functions are infinite-times continuously differentiable, this optimal interpolation theorem tells us the best approximation is a spline with "infinite" order. We have shown that the sampling function of B-spline interpolation with infinite order converge to the Shannon sampling functions, i.e., (5.9). Thus, we conclude that the optimal interpolation for limited sampling points from bandlimited signals is a Shannon sampling function, i.e., (5.8). Figure 5.2 shows some examples of these interpolations.

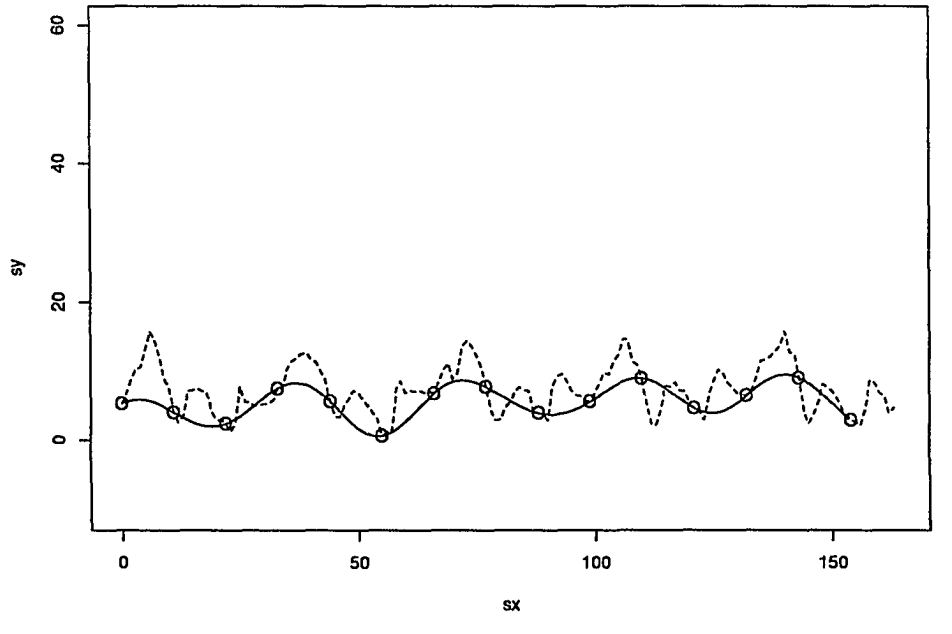


(a)

Figure 5.2: Cardinal series of Shannon Sampling (a) 15 sample points on a cross section of end milled surface. (b) 20 sample points on a cross section of bored surface. (c) 15 sample points on a cross section of shaped surface. Solid line represents the interpolated profile; dashed line represents the true surface profile (164 points).



(b)



(c)

Figure 5.2: (Continued)

CHAPTER 6. CASE STUDIES

The surface data used in this chapter are from the *Atlas of Machined Surfaces* (Stout, et al., 1990). We would like to thank Dr. Sullivan, P.J. for sending us these data. The 3-D logging system used to obtain this surface data is based on a modified Rank Taylor Hobson (RTH) Talysurf 5 surface measuring instrument incorporating a linear translation stage. The logged area for each surface were square with sides measuring 1.304 mm. The grid spacing on both axes was 8μm, giving a total of 26,896 (164$\times$ 164) data points for each sample. The material used in all machining processes was a free cutting mild steel. Appendix C shows perspective plots of five surface data. These five typical surfaces come from common manufacturing processes, namely, end milling (eml), grinding (sgl), fly cut (ftl), boring (bocll), and shaping (shlc). Each manufacturing process produces a surface with its own characteristic topography.

We use these five typical manufactured surfaces to compare the performance of different methods in evaluating the flatness errors under the uniform sampling situation. The methods are single points (straight line interpolation), universal kriging, and Shannon sampling series, i.e., (5.3) with $n = 2$. The derived random sampling results (described in Chapter 3) also provided a reference for the performance of these methods.

Band-limited surface profiles

In the previous chapter, we conjectured that the measured manufactured surface profiles were band-limited signals without any evidence. Two-dimensional spectral analysis provides a comprehensive description of both the structure and scales of pattern in a spatial sampling data set (Renshaw and Ford, 1983). We performed this technique on these five surfaces in order to determine if they are bandlimited signals. We used the two-dimensional Fast Fourier Transform program in the MATLAB (Etter, 1993) software package and obtained spectral plots for these five surfaces, which are shown in Appendix D. From these spectral plots, we see that the spectrum falls to zero within a finite range in both x and y axes for all of the five surfaces and the directions of travel of the waves can be seen clearly in the plots. Thus, we have demonstrated that the surface profiles are composed of a limited number of frequency components and show a significant pronounced lay and directional character of typical manufactured surfaces. These results concur with our previous conjecture.

Random sampling

The minimum zone mean profile (which is a plane when we consider the flatness error) is calculated from total of 26,896 points for each surface. The histograms of the deviations from minimum zone mean profile are shown in Figure 6.2. Table 6.1 shows the parameters of beta distribution for these deviations and the (true) flatness error for each surface. The mean and standard deviation of flatness errors estimation for a given number of sample points on each surface are calculated by (3.4) and (3.6). They are shown in Figure 6.6 with symbol "o" and "x" respectively.

Table 6.1: The parameters of beta distribution and flatness errors

Surface	a	b	α	β	Flatness error
bocl1	-20.5886	20.812	1.09043	3.03299	41.164028
eml	-9.10802	8.65788	5.04083	3.75953	17.300614
ftl	-3.86613	3.9746	4.09118	4.15834	7.68158
sgl	-1.95263	1.967854	1.52657	6.89763	4.402485
shlc	-8.01422	8.00488	2.0202	2.6777	15.983605

Universal kriging

As we discussed in Chapter 4, the empirical correlogram (variogram) is hard to identify and fit when we have small number of sampling points. If we take 1000 random sampling points from these five surfaces and draw the correlograms, Figure 6.3 show there are obvious wave correlation patterns between spatial locations for surface bocl1, eml, and shlc. Table 6.2 lists the parameters of exponential and wave models obtained by fitting the correlation functions by eye to these surfaces. In contrast to these clear correlation patterns drawn from large number of samples, the correlation patterns are impossible to identify for a small number of sample points, e.g., Figure 6.4 shows the correlograms of 100 and 25 uniform sampling points taken from bocl1 surface. The points shown in the correlogram plots are $n_d \geq 6$ pairs for a given distance d which is more relaxed than the recommended $n_d > 30$. Thus, we use the models listed in Table 6.2 as a priori correlation functions for universal kriging in this comparative study.

Table 6.2: The parameters of exponential and wave models

Surface	a_e (exponential model)	a_w (wave model)
bocl1	7.0	5.5
em1	5.0	4.4
ft1	10.5	N/A
sg1	2.5	N/A
sh1c	5.0	5.0

Shannon sampling

We use (5.3) with $n = 2$ as the interpolation function to approximate the five surfaces from a limited number of sampling points. Three samples of rectangular uniform sampling from different regions on the surface with 25, 49, and 100 points are taken from each surface respectively. The Nyquist frequency having a wavelength equal to two times the distance between two adjacent sample points is the highest frequency we can detect from the samples. For sample 1 of 25 points, the Nyquist frequency (W_1 and W_2) is $2\pi/(2*30) = \pi/30$ which is shown in Figure 6.1 (where 30 is the distance). Table 6.3 shows the x-y coordinates of these samples and W_1, W_2 for the formula (5.3). The Shannon sampling series (5.3) and universal kriging with the *a priori* correlation functions are used to approximate the real surfaces. The reconstructed surfaces are then discretized and the flatness errors are calculated from these discretized points. Figure 6.5 shows an example of the reconstructed surfaces by Shannon sampling series and universal kriging from 100 sample points on bocl1 surface. Figure 6.6 shows the overall results of flatness errors estimation by straight line interpolation, Shannon sampling series, universal kriging, and random sampling for the same set of 100 points. The results shown for universal kriging are based on wave models for bocl1, em1, sh1c, and exponential models for ft1, sg1 surface.

Table 6.3: Samples for the comparative study

Number of points /sample	Number of points for each xy axis	sample #	Coordinates of Bonding Box lower left corner coordinate	upper right corner coordinate	W_1, W_2
25	5	1	(30,30)	(150,150)	$\pi/30$
		2	(20,20)	(140,140)	$\pi/30$
		3	(15,15)	(159,159)	$\pi/36$
49	7	1	(3,3)	(159,159)	$\pi/26$
		2	(6,6)	(156,156)	$\pi/25$
		3	(9,9)	(153,153)	$\pi/24$
100	10	1	(1,1)	(163,163)	$\pi/18$
		2	(10,10)	(163,163)	$\pi/17$
		3	(5,5)	(158,158)	$\pi/17$

Summary

Based on this comparative study, we make the following observations:

1. From Figure 6.5, we see the performances of Shannon sampling series and universal kriging (with *a priori* correlation function) are quite similar. As we mentioned before, both methods are considered as "optimal" in stochastic prediction and for bandlimited functions approximation. Thus, we conjecture if the *a priori* correlation function can model the spatial dependence perfectly (the fitted models shown in Table 6.2 and Figure 6.3 are not rigorous fits), their results should agree with each other.

2. Figure 6.5(b) and 4.4(a) show large discrepancy of universal kriging results even we use *a priori* correlation function. The discrepancy may be due to the different sampling methods, i.e., uniform sampling and random sampling.

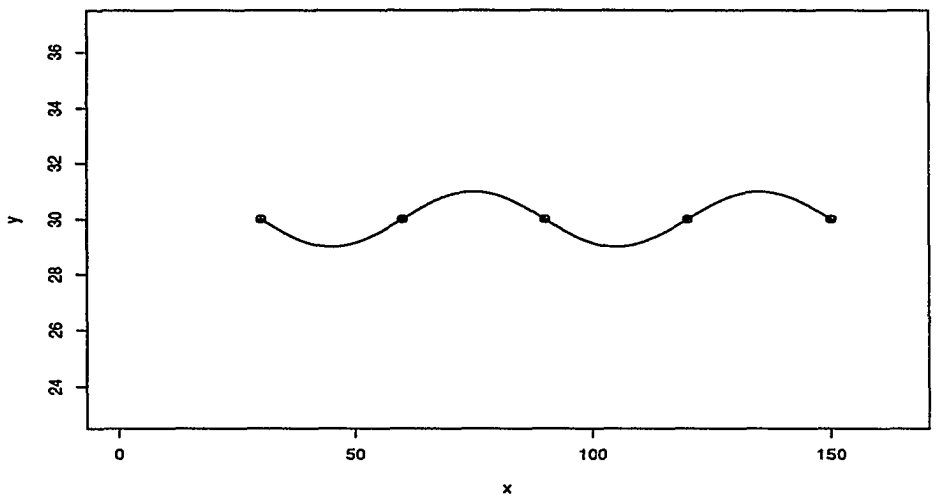


Figure 6.1: Nyquist frequency, $\pi/30$, for sample 1 of 25 points

3. Since many manufactured surfaces have pronounced lay and direction character (i.e., strong periodicity), the sampling period should avoid the surface periodicity as noted by Ripley (1981). If we take this into account when we perform uniform sampling with small number of samples, the result should be better (higher detectability and lower standard deviation) than random sampling, which has larger standard errors when a small number of samples is taken.

4. The form errors estimated by both interpolation methods are equal or greater than those calculated by single points. (Note: However, this is not the case for exponential models used by universal kriging. The krigged surfaces lie within the 3D convex hull which agrees with the observation we made in Chapter 4.) Therefore, if we use these optimal interpolation methods to estimate form errors, the type I error (reject the good part) should increase and type II error (accept a bad part) should decrease which is favorable especially when the part

will affect safety considerations in future use.

5. Kriging provides a standard deviation for the predicted surface. Spline interpolation also gives the error bound for the approximation, e.g., (5.10) and Kamada, et al. (1991). How to utilize this error information to characterize the mean and standard deviation of form error estimation for given number of uniform sampling points on general surfaces requires further study.

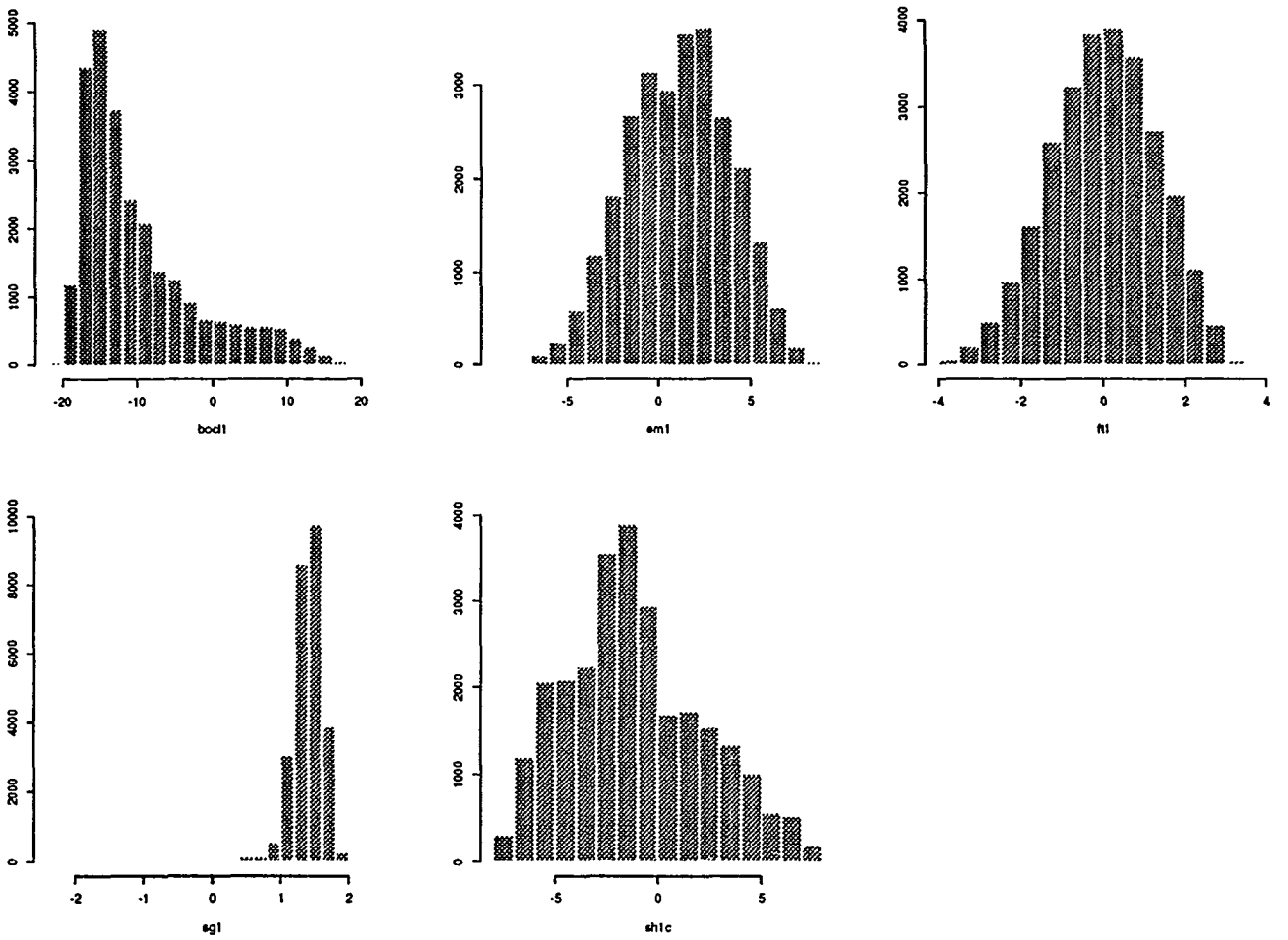


Figure 6.2: Histograms of deviations from minimum zone mean profiles

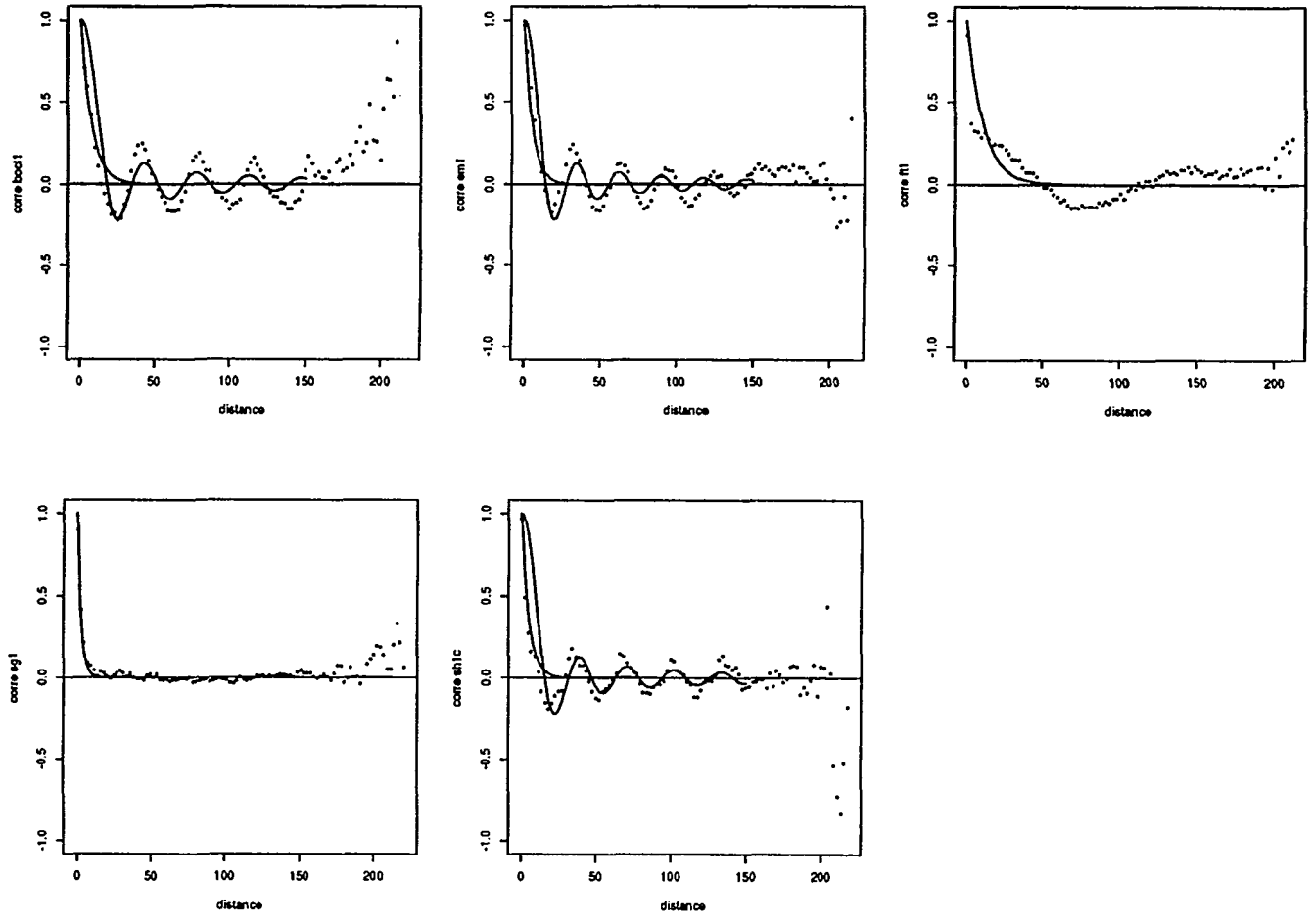


Figure 6.3: Correlograms obtained from 1000 points each surface

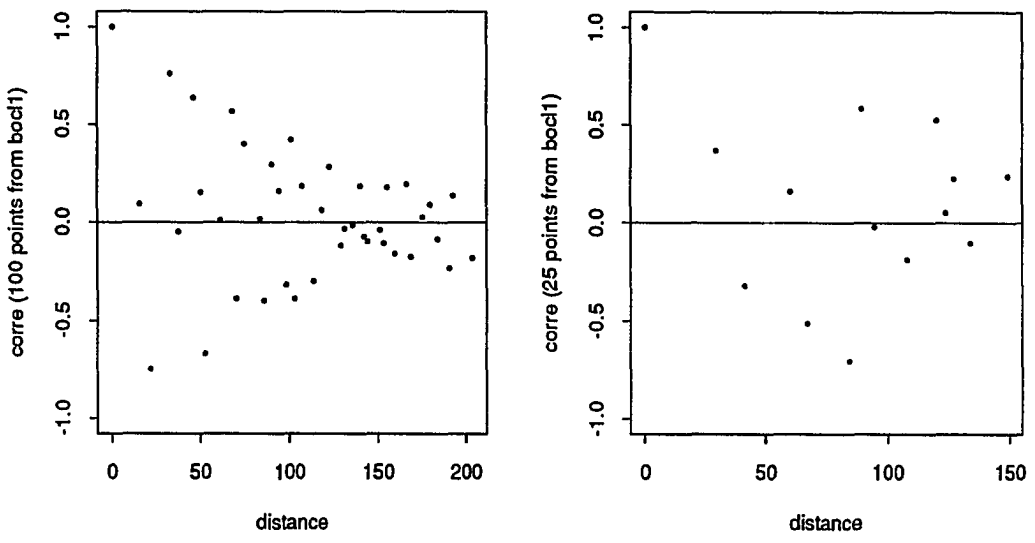
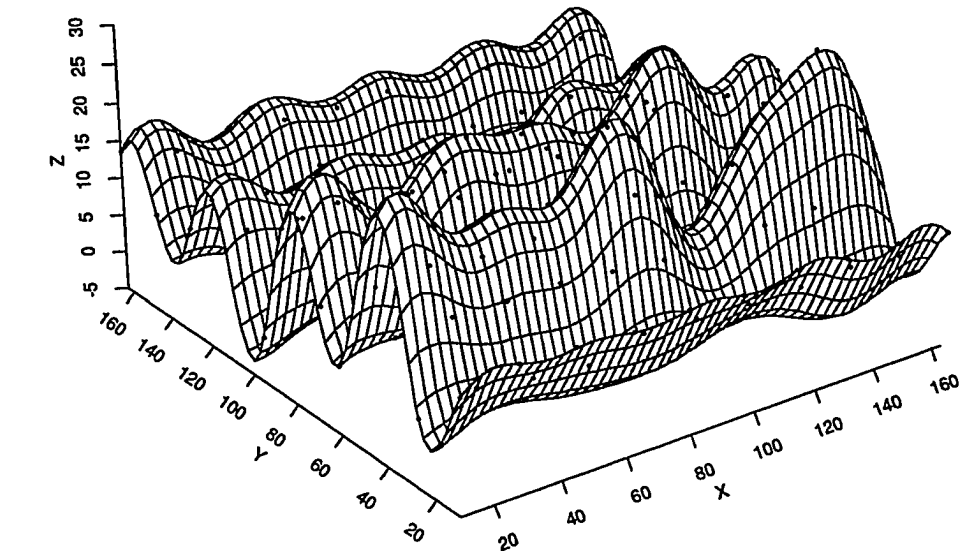
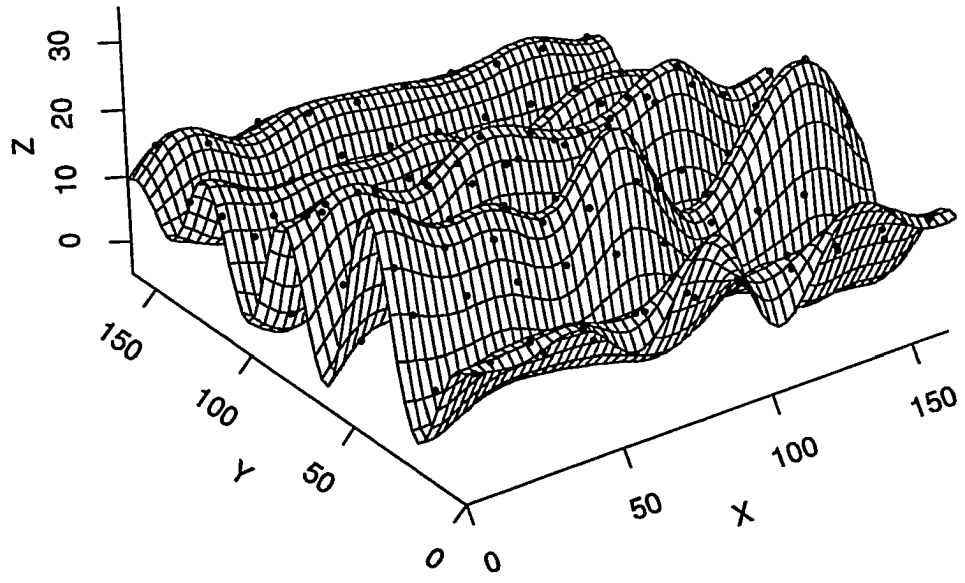


Figure 6.4: Correlograms of 100 and 25 points from bocl1 surface



(a)



(b)

Figure 6.5: (a) Cardinal series of Shannon Sampling in 3D, result of 100 sampling points taken from surface bocl1, (b) Universal kriging result of the same sampling points.

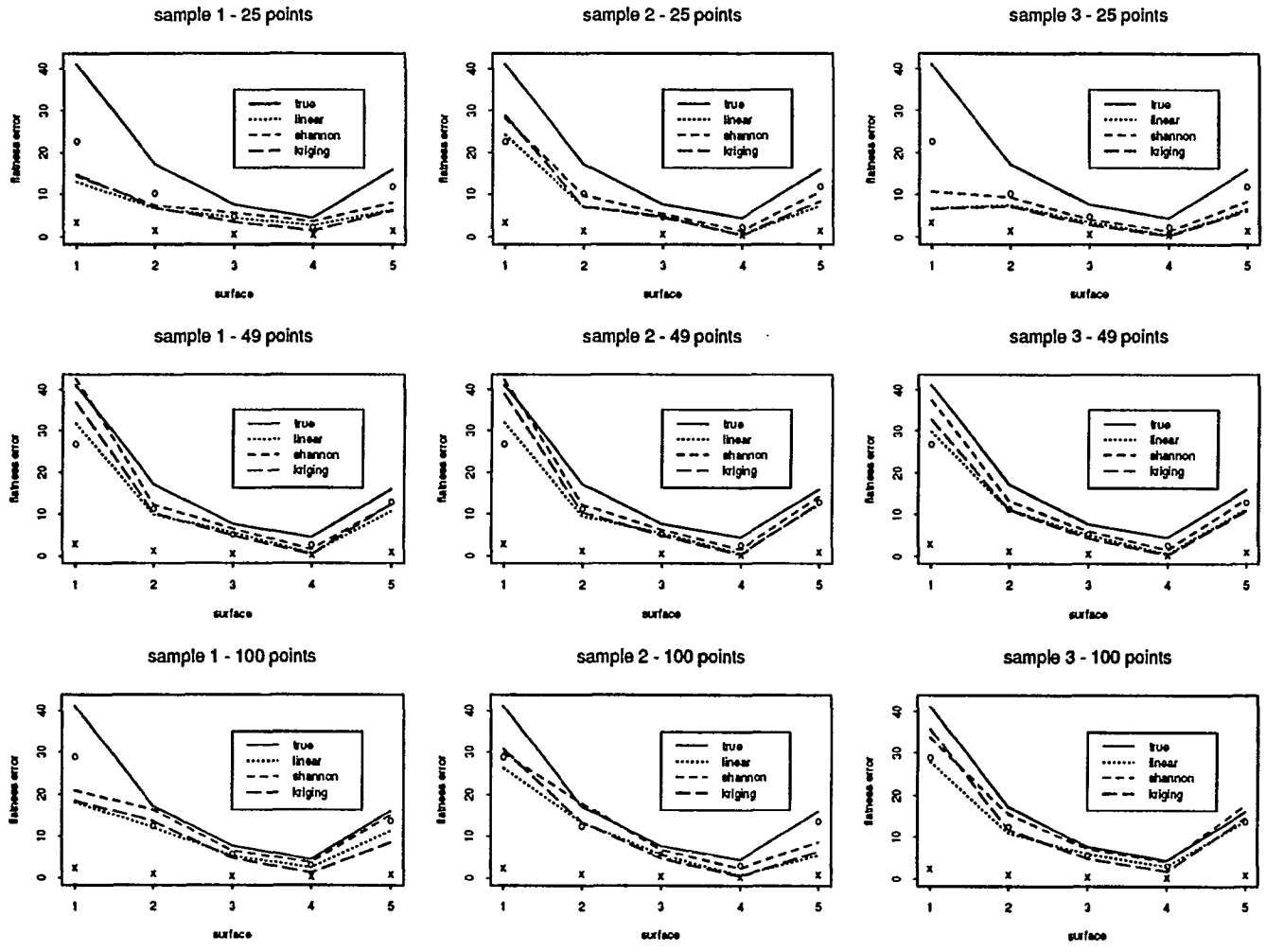


Figure 6.6: Overall results

CHAPTER 7. CONCLUSIONS

In a discussion at the 1993 International Forum on Dimensional Tolerancing and Metrology, the following points were made.

> In general, we see a point of diminishing returns, after which increasing the number of samples brings no advantage. However, we found that a plot of size vs. number of data points oscillates slightly as it converges, and certain numbers of samples lead to larger errors than adjacent numbers. (For example, 12 points might be worse than 11 or 13 points.) We don't know why this occurs, but it seems to be very repeatable for a given probe and machine. (p. 299)

> Our problem, as I see it, is the more points you take, the bigger the value of form error you get. So, we have a curve like this (trending upward as the number of points approached infinity. I don't know of any solid way of estimating, from somewhere here out to infinity, where that curve will go. (p. 301)

Caskey et al. (1992) and Hocken (1993) also reported the similar problems. In this thesis we have presented a probabilistic viewpoint of the problem by determining the theoretical form error distributions for a wide variety of surfaces profiles

under various points of inspection. Our results concur with empirical evidence of others and indicate that current practices in the evaluation of form are insufficient in dealing with the variety of surface profiles that one encounters due to different manufacturing processes and materials. The role of the surface profile distribution must be understood before accurate estimates of form error can be obtained.

Another issue raised in the curve fitting approach is the metric p selection for the fitting objective. From our studies, we have shown that the metric p selection does not appear to significantly affect the detectability. Sample size and the surface characteristics have the largest effect on detectability.

Kriging is often considered as an optimal interpolation procedure in the sense of correctly modeling the spatial dependence. Identifying and fitting a correct variogram model (covariance function) from a small number of sample points can be a difficult if not impossible task. Due to the lack of complete computer software because of uncertainty in the variogram estimation and the computational complexity, kriging does not have a significant advantage in the estimation of form error.

Finally, we applied the Shannon sampling theorem from communication engineering and have shown that the surface profiles are band-limited signals. We have shown also that the Shannon sampling function is in fact an infinite degree of B-spline interpolation function and thus a best approximation for band-limited signals. Both Shannon sampling series and universal kriging using *a priori* correlation functions were applied to the flatness error estimation for uniform sample points measured from five common engineering surfaces. The result shows both methods perform similarly. The probability of over-estimating form error increases and the probability of accepting bad parts decreases using interpolation methods versus using the points

directly.

Recommendations for further study

Most of the algorithms used to calculate the form errors shown in Appendix A are based on a curve-fitting (L_p-norm) approach. This approach can only obtain an approximate solution and possibly can not achieve global optimization. Deng (1993) gives criteria to achieve a global optimum for collected data in least squares fitting. More general conditions for general features is a subject of study. Hopp (1993) gives a start on the theory of testing metrology data analysis software. Several research issues still need to be addressed. The computational geometry approach to calculate the form errors is more difficult and computationally intensive but yields an exact solution for the sample points (but not exact with regard to form error). The solutions to many geometries are still unknown, e.g., finding the minimax approximating line (curve) of a set of points in 3-space. NURBS (NonUniform Rational B-Spline) curves and surfaces are commonly used in CAD systems. However, the inspection and analysis of these geometries has not been addressed.

Voelcker (1993) comments on the future of metrology, "CMM-based coordinate metrology is severely data-limited at present, because data are expensive when collected sequentially by moving machinery. It may be possible to finesse the currently vexing sample-set-sufficiency problem by moving to a data-rich environment based on wave phenomena rather than contact sensing. Thus the future of measurement per se is more likely to be paced by technology than theory." If the high speed data collection technology is reliable and available, the sampling strategies and sample data analysis techniques are another issue need to be addressed.

BIBLIOGRAPHY

[1] ANSI Y14.5M, 1982, Dimensioning and Tolerancing, American National Standard, Engineering Drawings and Related Documentation Practices, The American Society for Mechanical Engineers, New York.

[2] Akima, H., 1978, A Method of Bivariate Interpolation and Smooth Surface Fitting for Irregularly Distributed Data Points. *ACM Transactions on Mathematical Software*, 4, pp. 148–164.

[3] Anderson, D.H. and Osborne, M.R., 1977, "Discrete, Nonlinear Approximation Problems in Polyhedral Norms. A Levenberg-like Algorithm", *Numerical Mathematics*, 28, pp. 157–170.

[4] Barber, C.B., Dobkin, D.P., and Huhdanpaa, 1993, *The Quickhull Algorithm for Convex Hull*, Technical Report GCG53, The National Science and Technology Research Center for Computation and Visualization of Geometric Structures, University of Minnesota.

[5] Barrodale and Phillips, 1975, "Solution of an Overdetermined System of Linear Equation in the Chebychev Norm", *ACM Transactions on Mathematical Software*, 1, pp. 264–270.

[6] Becker, R. A., 1988, *The New S Language : A Programming Environment for Data Analysis and Graphics*, Pacific Grove, California: Wadsworth & Brooks/Cole Advanced Books & Software.

[7] Caskey, G., et al., 1991, Sampling Techniques for Coordinate Measuring Machines, *Proceedings of the 1991 NSF Design and Manufacturing Systems Conference*, pp. 779–786.

[8] Caskey, G., et al., 1992, Sampling Techniques for Coordinate Measuring Machines, *Proceedings of the 1991 NSF Design and Manufacturing Systems Conference*, pp. 983–988.

[9] Cressie, N., 1993, *Statistics for Spatial Data*, Revised Edition, New York: J. Wiley.

[10] Deng, J.J., 1993, *Criteria for Collected Data in Least Squares Circle Fitting*, PhD Dissertation, Department of Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa.

[11] Dhanish, P.B. and Shunmugam, M.S., 1991, An Algorithm for Form Error Evaluation — Using the Theory of Discrete and Linear Chebyshev Approximation, *Computer Methods in Applied Mechanics and Engineering*, 92, pp. 309–324.

[12] Etesami, F. and Qiao, H., 1989, Analysis of Two-dimensional Measurement Data For Automated Inspection, *Journal of Manufacturing Systems*, Vol. 9, No. 1, pp. 21–34.

[13] Etter, D. M., 1993, *Engineering Problem Solving with MATLAB*, Englewood Cliffs, N.J.: Prentice Hall.

[14] Gonin, R. and Money, A.H., 1989, *Nonlinear Lp-norm Estimation*, New York: Marcel Dekker.

[15] He, J.R., 1991, Estimating the Distributions of Manufactured Dimensions with Beta Probability Density Function, *International Journal of Machine Tool Design and Research*, Vol. 31, No. 3, pp. 383–396.

[16] Hocken, R.J., Raja, J., and Babu, U., 1993, "Sampling Issues in Coordinate Metrology", *Proceedings of the 1993 International Forum on Dimensional Tolerancing and Metrology*, edited by Srinivasan, V. and Voelcker, H.B., CRTD-Vol. 27, pp. 97–111.

[17] Hopp, T.H., 1993, Computational Metrology, *Proceedings of the 1993 International Forum on Dimensional Tolerancing and Metrology*, edited by Srinivasan, V. and Voelcker, H.B., CRTD-Vol. 27, pp. 207–217.

[18] Houle, M.E. and Toussaint, G.T., 1988, Computing the Width of a Set, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 5, pp. 761–765.

[19] Jerri, A.J., 1977, The Shannon Sampling Theorem — Its Various Extensions and Applications: A Tutorial Review, *Proceedings of the IEEE*, Vol. 65, No. 11, pp. 1565–1596.

[20] Kamada, M., Toraichi, K., and Ikebe, Y., 1991, A Note on Error Estimation for Spline Interpolation Method with Sampling Bases, *Electronics and Communications in Japan*, Part 3, Vol. 74, No. 4, pp. 51–59.

[21] Kendall, M. and Stuart, A., 1977, *The Advanced Theory of Statistics, Vol. 1, Distribution Theory*, Fourth edition, London & High Wycombe: Charles Griffin & Company Limited.

[22] Marks, R.J., 1991, *Introduction to Shannon Sampling and Interpolation*, New York: Springer-Verlag.

[23] Menq, C.H., Yau, H.T., and Lai, G.Y., 1990, Statistical Evaluation of Form Tolerances Using Discrete Measurement Data. *Proceedings of the Symposium on Advances in Integrated Product Design and Manufacturing*, the 1990 ASME Winter Annual Meeting, November 25–30, Dallas, TX, pp. 135–150.

[24] Murthy, T.S.R. and Abdin, S.Z., 1980, Minimum Zone Evaluation of Surfaces, *International Journal of Machine Tool Design and Research*, Vol. 20, No. 3, pp. 123–136.

[25] Murthy, T.S.R., 1982, A Comparison of Different Algorithms for Cylindricity Evaluation, *International Journal of Machine Tool Design and Research*, Vol. 22, No. 4, pp. 283–292.

[26] Nayak, P.R., 1971, Random Process Model of Rough Surfaces, *Transactions of ASME, Journal of Lubrication Technology*, Vol. 93F, pp. 398–407.

[27] Powell, 1981, *Approximation Theory and Methods,* Cambridge: Cambridge University Press.

[28] Ripley, B. D., 1981, *Spatial Statistics*, New York: J. Wiley.

[29] Roy, U. and Zhang, X., 1992, Establishment of a Pair of Concentric Circles with the Minimum Radial Separation for Assessing Roundness Error, *Computer Aided Design*, Vol. 24, No. 3, pp. 161–168.

[30] Sayles, R.S. and Thomas, T.R., 1977, The Spatial Representation of Surface Roughness by Means of the Structure Function: a Practical Alternative to Correlation, *Wear*, 42, pp. 263–276.

[31] Sayles, R.S. and Thomas, T.R., 1978, Surface Topography as a Nonstationary Random Process, *Nature*, 271, pp. 431–434.

[32] Sayles, R.S. and Thomas, T.R., 1979, Measurements of the Statistical Microgeometry of Engineering Surfaces, *Transaction of American Society of Mechanical Engineers. Journal of Lubrication Technology*, 101F, pp. 409–418.

[33] Shunmugam, M.S., 1986, On Assessment of Geometric Errors, *International Journal of Production Research*, Vol. 24, No. 2, pp. 413–425.

[34] Shunmugam, M.S., 1987, New Approach for Evaluating Form Errors of Engineering Surfaces, *Computer-Aided Design*, Vol. 19, No. 7, pp. 368–374.

[35] Shunmugam, M.S., 1987, Comparison of Linear and Normal Deviations of Forms of Engineering Surfaces, *Precision Engineering*, Vol. 9, No. 2, pp. 96–102.

[36] Shunmugam, M.S., 1990, Establishing Reference Figures for Form Evaluation of Engineering Surfaces, *Journal of Manufacturing Systems*, Vol. 10, No. 4, pp. 314–321.

[37] Stout, Davis and Sullivan, 1990, *Atlas of Machined Surfaces*, London: Chapman and Hall.

[38] Traband, M.T., Joshi,S., Wysk, R.A., Cavalier, T.M., 1989, Evaluation of Straightness and Flatness Tolerances Using the Minimum Zone, *Manufacturing Review*, Vol.2, No.3, pp.189–195.

[39] Unser, M., Aldroubi, A., Eden, M., 1993, B-spline Signal Processing, IEEE Transactions on Signal Processing, Vol. 41, No. 2, pp. 821–848.

[40] Venables, W.N. and Ripley, B.D., 1994, *Modern Applied Statistics with S-Plus*, New York: Springer.

[41] Walker, R.K. and Srinivasan, V., 1993, Creation and Evolution of the ASME Y14.5.1M Standard", *Proceedings of the 1993 International Forum on Dimensional Tolerancing and Metrology*, edited by Srinivasan, V. and Voelcker, H.B., CRTD-Vol. 27, pp. 19–30.

[42] Watson, D.F., 1992, *Contouring: A Guide to the Analysis and Display of Spatial Data*, New York: Pergamon Press.

[43] Woo, T., Liang, R., and S. Pollock, 1993, Hammersley Sampling for Efficient Surface Coordinate Measurements, *Proceedings of the 1993 NSF Design and Manufacturing Systems Conference*, pp. 1489–1495.

[44] Zayed, A.I., 1993, *Advances in Shannon's Sampling Theory*, Boca Raton: CRC Press.

APPENDIX A. FORM ERROR EVALUATION PROGRAMS

Table A.1: Form error evalution programs (I)

Form error	Method	Program
Straightness	L_2 (linear deviation)	straight.l2(file) (Splus lsfit function)
	L_2 (normal deviation)	straight.nl2(file) (Solve $\sum \hat{e}_i^2$ analytically)
		mlstraightnl2.c (MATLAB optimization toolbox: leastsq)
	L_∞ (normal deviation)	straight.li(file,η_1,η_2) (Osborne and Watson algorithm, see Gonin and Money (1989))
		mlstraightli.c (MATLAB optimization toolbox: minimax)
	Convex hull	straight.ch(file) (Splus chull function to obtain convex hull, schwidth.c to caluculate the minimum width of convex hull, algorithm see Houle and Toussaint (1988))

Table A.2: Form error evalution programs (II)

Form error	Method	Program
Flatness	L_2 (linear deviation)	flat.l2(file) (Splus lsfit function)
	L_2 (normal deviation)	flat.nl2(file) (Splus ms function)
		mlflatnl2.c (MATLAB optimization toolbox: leastsq)
	L_∞ (normal deviation)	flat.li(file,η_1,η_2,η_3) (Osborne and Watson algorithm, see Gonin and Money (1989))
		mlflatli.c (MATLAB optimization toolbox: minimax)
Circularity	L_2 (normal deviation)	cir.nl2(file) (Splus ms function)
		mlcirnl2.c (MATLAB optimization toolbox: leastsq)
	L_∞ (normal deviation)	cir.li(file,η_1,η_2,η_3) (Osborne and Watson algorithm, see Gonin and Money (1989))
		mlcirli.c (MATLAB optimization toolbox: minimax)
Cylindricity	L_∞ (linear deviation)	cylindricity file η_1 η_2 η_3 η_4 η_5 (Osborne and Watson algorithm, see Gonin and Money (1989))

APPENDIX B. RESULTS COMPARISON OF PUBLISHED DATA SETS

1. Straightness $y = \eta_1 + \eta_2 x$

Data set 1 (Shunmugam (1986, 1987ab, 1990), Traband et al. (1989), Dhanish and Shunmugam (1991))

x	−2	−1	0	1	2
y	3	5	2	1	2

Method	η_1	η_2	Form error
L_2 (LD)	2.6	−0.6	2.8
L_2 (ND)	2.6	−0.9355531	2.143676
L_∞ (ND)	2.500001	−0.9999995	2.121321

Note: LD=Linear Deviation; ND=Normal Deviation

Data set 2 (Traband et al.(1989))

x	1	2	3	4	5
y	2.428	2.891	3.445	2.931	3.895
x	6	7	8	9	10
y	4.196	4.497	4.662	4.545	4.303

Method	η_1	η_2	Form error
L_2 (LD)	2.4614	0.2396182	0.9128545
L_2 (ND)	2.445333	0.2425394	0.8956511
L_∞ (ND)	2.456333	0.2286667	0.8578577

Data set 3 (Traband et al. (1989))

x	y	x	y
0.05	-0.066450	0.55	-0.095250
0.10	-0.064380	0.60	-0.011540
0.15	0.008761	0.65	-0.024060
0.20	-0.011170	0.70	0.035150
0.25	0.062370	0.75	-0.019970
0.30	-0.038290	0.80	0.015400
0.35	0.065500	0.85	-0.013240
0.40	0.063570	0.90	-0.022250
0.45	0.028490	0.95	0.077100
0.50	-0.006113	1.00	-0.000360

Method	η_1	η_2	Form error
L_2 (LD)	-0.01628805	0.02943136	0.1666363
L_2 (ND)	-0.01666975	0.03015843	0.1667059
L_∞ (ND)	-0.023575	0.01933333	0.1645859

Data set 4 (Traband et al. (1989))

x	y	x	y
0.3952	-0.0032	2.6001	0.0007
0.6953	-0.0016	2.8590	0.0017
0.9669	-0.0042	3.0662	0.0025
1.2762	-0.0028	3.2165	-0.0017
1.5797	-0.0037	3.4217	0.0026
1.8593	-0.0007	3.6179	0.0027
2.1333	-0.0010	3.8185	0.0047
2.4197	0.0007		

Method	η_1	η_2	Form error
L_2 (LD)	-0.004953762	0.002093011	0.005376898
L_2 (ND)	-0.00495377	0.002093014	0.005376895
L_∞ (ND)	-0.005595367	0.002017162	0.005185658

Data set 5 (Traband et al. (1989))

x	y
0.2845	−0.0034
0.6600	−0.0032
1.2041	−0.0030
1.4994	−0.0035
1.8494	−0.0036
2.2261	−0.0025
2.5724	−0.0028
2.9076	−0.0026
3.2548	−0.0031
3.4142	−0.0031
3.6307	−0.0029
3.9237	−0.0029
4.2647	−0.0028
4.5112	−0.0028
4.8150	−0.0027
5.1334	−0.0027
5.3603	−0.0030
5.6534	−0.0032
5.9058	−0.0020
6.0774	−0.0019
6.2962	−0.0019
6.5240	−0.0019
6.7114	−0.0017
6.9996	−0.0019
7.2076	−0.0017

Method	η_1	η_2	Form error
L_2 (LD)	−0.003588502	0.0002226964	0.001463247
L_2 (ND)	−0.003588502	0.0002226964	0.001463247
L_∞ (ND)	−0.003552695	0.0001783604	0.001311295

2. Flatness $z_i = \eta_1 + \eta_2\, x + \eta_3\, y$

Data set 1 (Shunmugam (1986, 1987ab, 1990), Traband et al. (1989), Dhanish and Shunmugam (1991))

x	y	z
−2	1	5
−1	1	4
0	1	1
1	1	2
2	1	2
−2	0	4
−1	0	3
0	0	3
1	0	2
2	0	2
−2	−1	3
−1	−1	4
0	−1	2
1	−1	1
2	−1	2

Method	η_1	η_2	η_3	Form error
L_2 (LD)	2.666667	−0.6	0.2	2.8
L_2 (ND)	2.666666	−0.729863	0.4289852	2.532129
L_∞ (ND)	2.50	−0.75	−0.25	1.961161

Minimum zone = 2.0000 (Traband et al., 1989)

Data set 2 (Murthy and Abdin (1980), Traband et al. (1989))

x	y	z
0.00	0.00	2.00
0.00	25.00	5.00
0.00	50.00	6.00
0.00	75.00	8.00
0.00	100.00	9.00
25.00	0.00	5.00
25.00	25.00	7.00
25.00	50.00	8.00
25.00	75.00	9.00
25.00	100.00	12.00
50.00	0.00	6.00
50.00	25.00	7.00
50.00	50.00	8.00
50.00	75.00	9.00
50.00	100.00	11.00
75.00	0.00	7.00
75.00	25.00	7.00
75.00	50.00	6.00
75.00	75.00	7.00
75.00	100.00	9.00
100.00	0.00	7.00
100.00	25.00	6.00
100.00	50.00	6.00
100.00	75.00	6.00
100.00	100.00	8.00

Method	η_1	η_2	η_3	Form error
L_2 (LD)	5.16	0.00080	0.0408	5.9
L_2 (ND)	5.156904	0.0008011909	0.04086073	5.888982
L_∞ (ND)	4.431818182	0.001818182	0.050909091	4.857338

Minimum zone = 6.2343 (Traband et al., 1989)

Data set 3 (Traband et al. (1989))

x	y	z
0.20	0.20	−0.066450
0.20	0.40	−0.064380
0.20	0.60	0.008761
0.20	0.80	−0.011170
0.20	1.00	−0.062370
0.40	0.20	−0.038290
0.40	0.40	0.065500
0.40	0.60	0.063570
0.40	0.80	0.028490
0.40	1.00	−0.006113
0.60	0.20	−0.095250
0.60	0.40	−0.011540
0.60	0.60	−0.024060
0.60	0.80	0.035150
0.60	1.00	−0.019970
0.80	0.20	0.015400
0.80	0.40	−0.013240
0.80	0.60	−0.022250
0.80	0.80	0.077100
0.80	1.00	−0.000360
1.00	0.20	0.057730
1.00	0.40	−0.056200
1.00	0.60	0.092060
1.00	0.80	0.065360
1.00	1.00	−0.021210

Method	η_1	η_2	η_3	Form error
L_2 (LD)	−0.05526864	0.06101914	0.03084648	0.1667845
L_2 (ND)	−0.0567301	0.06263704	0.03166436	0.1665349
L_∞ (ND)	−0.044235	0.026200	0.054200	0.154870

Minimum zone = 0.1756 (Traband et al., 1989)

Data set 4 (Traband et al. (1989))

x	y	z
0.3846	0.2416	-0.0828
1.5008	0.2922	-0.0821
2.3107	0.3289	-0.0787
2.9817	0.3593	-0.0789
3.6964	0.3917	-0.0760
3.6743	0.8794	-0.0785
3.1195	0.8543	-0.0735
2.3552	0.8196	-0.0745
1.5875	0.7849	-0.0714
0.5573	0.7382	-0.0740
0.5413	1.0921	-0.0730
1.2205	1.1229	-0.0727
2.1673	1.1658	-0.0716
3.0881	1.2076	-0.0749
3.8459	1.2419	-0.0799
3.8305	1.5796	-0.0848
3.2057	1.5514	-0.0410
2.4230	1.5159	-0.0759
1.6710	1.4819	-0.0746
0.5263	1.4300	-0.0745

Method	η_1	η_2	η_3	Form error
L_2 (LD)	-0.08152902	-0.000073793	0.007368128	0.04396168
L_2 (ND)	-0.08153099	-0.000073940	0.007370543	0.04396046
L_∞ (ND)	-0.091755442	-0.004399395	0.028318718	0.04183267

Minimum zone = 0.04185 (Traband et al., 1989)

Data set 5 (Traband et al. (1989))

x	y	z
0.2556	0.2994	0.0005
1.4992	0.3371	0.0013
2.6656	0.3726	0.0000
3.5978	0.4009	0.0005
4.6241	0.4321	-0.0007
4.5989	1.2640	0.0001
3.4451	1.2289	0.0008
2.7096	1.2066	0.0004
1.6726	1.2968	0.0014
0.5273	1.2620	0.0009
0.1683	2.1414	-0.0002
0.9906	2.1663	0.0010
2.5485	2.1801	0.0008
3.4605	2.1369	0.0011
4.8632	2.1795	-0.0017
4.8401	2.9417	-0.0014
3.6557	2.9058	0.0012
2.4224	2.8683	0.0012
1.3839	2.8368	0.0011
0.4966	2.8098	-0.0002
0.4672	3.7751	-0.0008
1.6709	3.8116	0.0010
2.8864	3.8486	0.0006
3.7562	3.8750	0.0008
4.6746	3.9029	-0.0003

Method	η_1	η_2	η_3	Form error
L_2 (LD)	0.0009525604	-0.000186619	-0.000047496	0.002709154
L_2 (ND)	0.0009525606	-0.000186619	-0.000047496	0.002709154
L_∞ (ND)	0.0002603628	-0.000172624	0.0000884590	0.002627309

Minimum zone = 0.002817 (Traband et al., 1989)

3. Circularity

Data set 1: (Shunmugam (1986))

(1) $r_i = \eta_1 + \eta_2\ cos(\phi_i) + \eta_3\ sin(\phi_i)$

ϕ_i	r_i	LS residuals	L_∞ residuals
0	4	-0.8535534	-1.1213203
45	4	-0.0428932	-0.2928932
90	3	1.2071068	1.1213203
135	5	-1.2500000	-1.1213203
180	2	0.8535534	1.1213203
225	3	-0.9571068	-0.7071068
270	1	0.7928932	0.8786797
315	2	0.2500000	0.1213203
	h_t	2.4571068	2.2426406
	η_1	3.0000000	3.0000000
	η_2	0.1464466	-0.1213203
	η_3	1.2071068	1.1213203

The deviation is 2.2433 as reported in Roy, U. and Zhang, X., 1992, Establishment of a pair of concentric circles with the minimum radial separation for assessing roundness error, *Computer Aided Design*, Vol. 24, No. 3, pp. 161–168.

(2) $(x_i - \eta_1)^2 + (y_i - \eta_2)^2 = \eta_3^2$

x	y	LS residuals	L_∞ residuals
4.0000000	0.0000000	0.9495872	0.9264926
2.8284271	2.8284271	0.0093216	0.1552449
0.0000000	3.0000000	-1.4801608	-1.3168030
-3.5355339	3.5355339	0.9645387	0.9264926
-2.0000000	0.0000000	-0.9037659	-1.1985311
-2.1213203	-2.1213203	0.7190288	0.3826191
0.0000000	-1.0000000	-1.0113835	-1.3168030
1.4142136	-1.4142136	-0.2387083	-0.4561400
	h_t	2.4446995	2.2432956
	η_1	-0.0096779	-0.1279891
	η_2	1.2343914	1.0000000
	η_3	3.2457959	3.3208941

Data set 2:

(1) $r_i = \eta_1 + \eta_2\ cos(\phi_i) + \eta_3\ sin(\phi_i)$

ϕ_i	r_i	LS residuals	L_∞ residuals
0.000000	1.037000	-0.029637	-0.021569
36.000000	0.992320	0.014865	0.021569
72.000000	0.999941	0.002942	0.004700
108.000000	0.994568	0.001534	-0.003348
144.000000	0.985114	0.004318	-0.006361
180.000000	0.992000	-0.006580	-0.019998
216.000000	0.995113	-0.009514	-0.021569
252.000000	0.989503	0.000397	-0.006711
288.000000	0.974644	0.022038	0.021569
324.000000	1.003715	-0.000364	0.004964
	h_t	0.051674946	0.04313782
	η_1	0.9963918	0.9937163
	η_2	0.0109714	0.0217148
	η_3	0.0032601	0.0044310

(2) $(x_i - \eta_1)^2 + (y_i - \eta_2)^2 = \eta_3^2$

x	y	LS residuals	L_∞ residuals	
1.0370	0.0000	0.0291323	0.0212738	0.021507
0.8030	0.5830	-0.0152233	-0.0217392	-0.021507
0.3090	0.9510	-0.0030227	-0.0046536	-0.004421
-0.3070	0.9460	-0.0013594	0.0034576	0.003690
-0.7970	0.5790	-0.0040147	0.0063303	0.006563
-0.9920	0.0000	0.0068488	0.0197631	0.019996
-0.8050	-0.5850	0.0096426	0.0212738	0.021507
-0.3060	-0.9410	-0.0004620	0.0064981	0.006731
0.3010	-0.9270	-0.0223191	-0.0217392	-0.021507
0.8120	-0.5900	-0.0000984	-0.0052153	-0.004983
	h_t	0.0514514	0.0430131	0.0430131
	η_1	0.0113582	0.0217446	0.021745
	η_2	0.0031419	0.0042836	.004284
	η_3	0.9965143	0.9939906	.993785

4. Sphericity

$r_i = \eta_1 + \eta_2\, cos(\beta_i)\, cos(\phi_i) + \eta_3\, cos(\beta_i)\, sin(\phi_i) + \eta_4\, sin(\beta_i)$

data set (Dhanish and Shunmugam (1991))

ϕ_i	β_i	r_i	LS residuals	L_∞ residuals
0	90	5	1.65947	1.41418
0	45	5	1.28600	0.58578
0	0	4	0.39837	-0.41421
0	-45	3	-0.06922	-0.58581
0	-90	3	0.57134	0.58575
45	45	3	-0.98648	-1.24265
45	0	4	0.01304	-0.17158
45	-45	2	-1.34170	-1.41424
90	45	4	0.19766	0.41419
90	0	3	-0.72655	-0.24268
90	-45	4	0.84244	1.24261
135	45	3	-0.26945	0.17155
135	0	3	0.02706	0.82838
135	-45	3	0.37533	0.99997
180	45	1	-1.69996	-1.41420
180	0	3	0.83243	1.41420
180	-45	2	-0.05518	0.41421
225	0	2	0.21781	0.17163
225	-45	2	0.21733	0.24268
225	-90	3	0.57139	0.58582
270	45	2	-0.61156	-1.24255
270	0	2	-0.04256	-0.75725
270	-45	1	-0.96678	-1.41414
315	45	3	-0.14444	-0.99993
315	0	3	0.20385	-0.82832
315	-45	2	-0.49966	-1.17151
		h_t	3.359429	2.8284395
		η_1	2.884574	3
		η_2	0.717052	1.414214
		η_3	0.841958	0.242641
		η_4	0.455935	0.585786

5. Cylindricity

$r_i = \eta_1 + \eta_2 cos(\phi_i) + \eta_3 sin(\phi_i) + \eta_4\ z_i\ cos(\phi_i) + \eta_5\ z_i\ sin(\phi_i)$

data set (Dhanish and Shunmugam (1991))

ϕ_i	z_i	r_i	LS residuals	L_∞ residuals
0	1	5	1.11485	0.95119
0	0	4	0.43807	0.14645
0	−1	3	−0.23871	−0.65829
45	1	3	−1.01394	−0.79637
45	0	4	0.17800	0.12622
45	−1	4	0.36993	0.04882
90	1	4	0.57324	0.95116
90	0	3	−0.47854	−0.35358
90	−1	3	−0.53032	−0.65831
135	1	3	0.53243	0.75589
135	0	3	0.26726	0.40233
135	−1	2	−0.99790	−0.95122
180	1	1	−0.69823	−0.85356
180	0	3	0.97855	0.95116
180	−1	3	0.65532	0.75589
225	1	2	0.43061	−0.10596
225	0	2	0.23866	−0.02857
225	−1	1	−0.95329	−0.95118
270	1	2	−0.15653	−0.85347
270	0	2	−0.10477	−0.54874
270	−1	2	−0.05300	−0.24402
315	1	3	−0.11571	−0.65821
315	0	3	0.14944	−0.30465
315	−1	2	−0.58540	−0.95109
		h_t	2.12879	1.90241
		η_1	2.791674	2.951184
		η_2	0.770257	0.902369
		η_3	0.686848	0.402369
		η_4	0.32322	0.195262
		η_5	−0.051791	−0.304738

APPENDIX C. MACHINED SURFACES

This appendix shows five perspective plots of machined surfaces. The manufacturing processes are boring process, end milling process, fly cut process, fine grounding process, and shaping process. Data are from *Atlas of Machined Surfaces* (Stout, et al., 1990). We would like to thank Dr. Sullivan, P.J. for sending us these data.

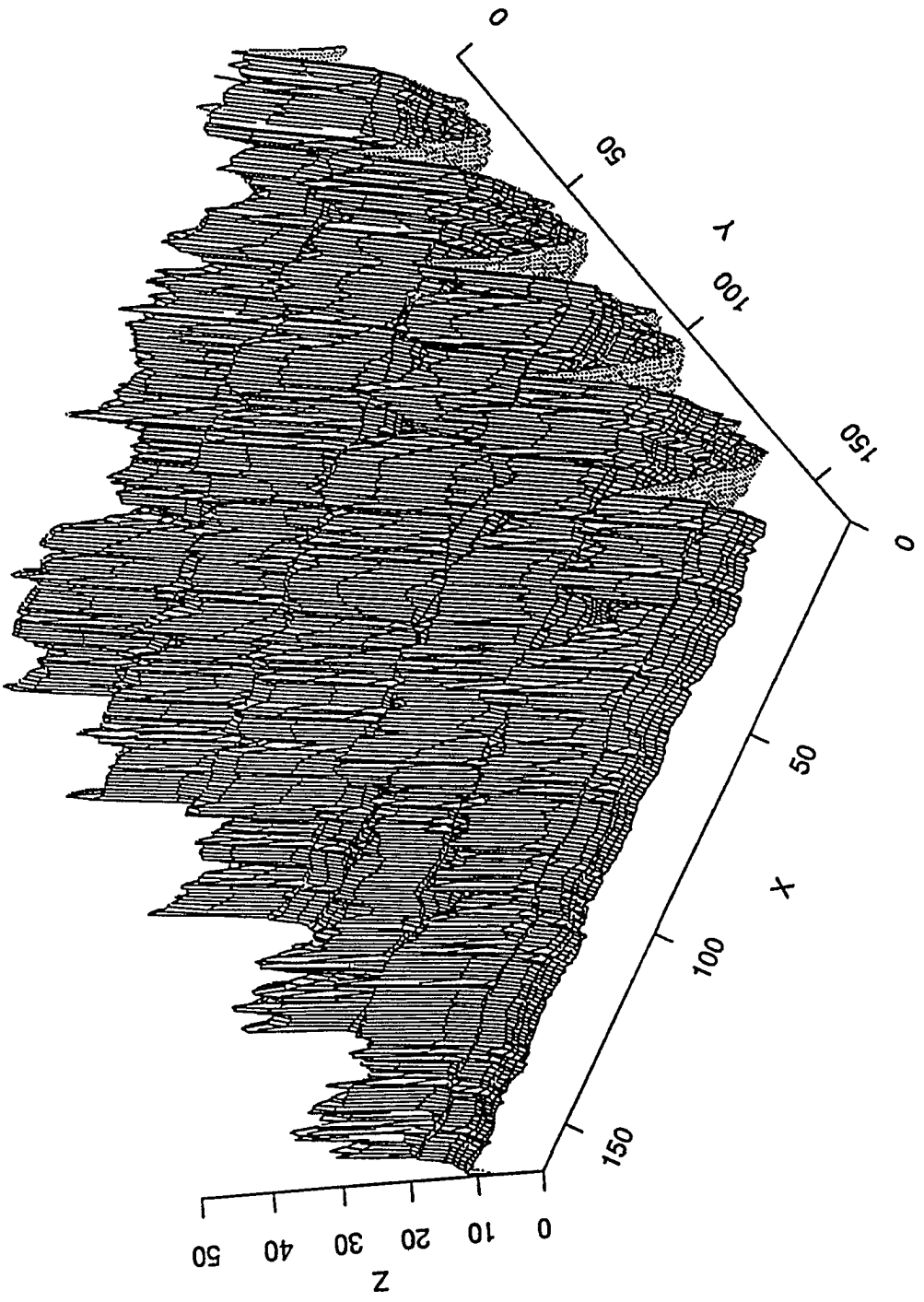


Figure C.1: Bored surface

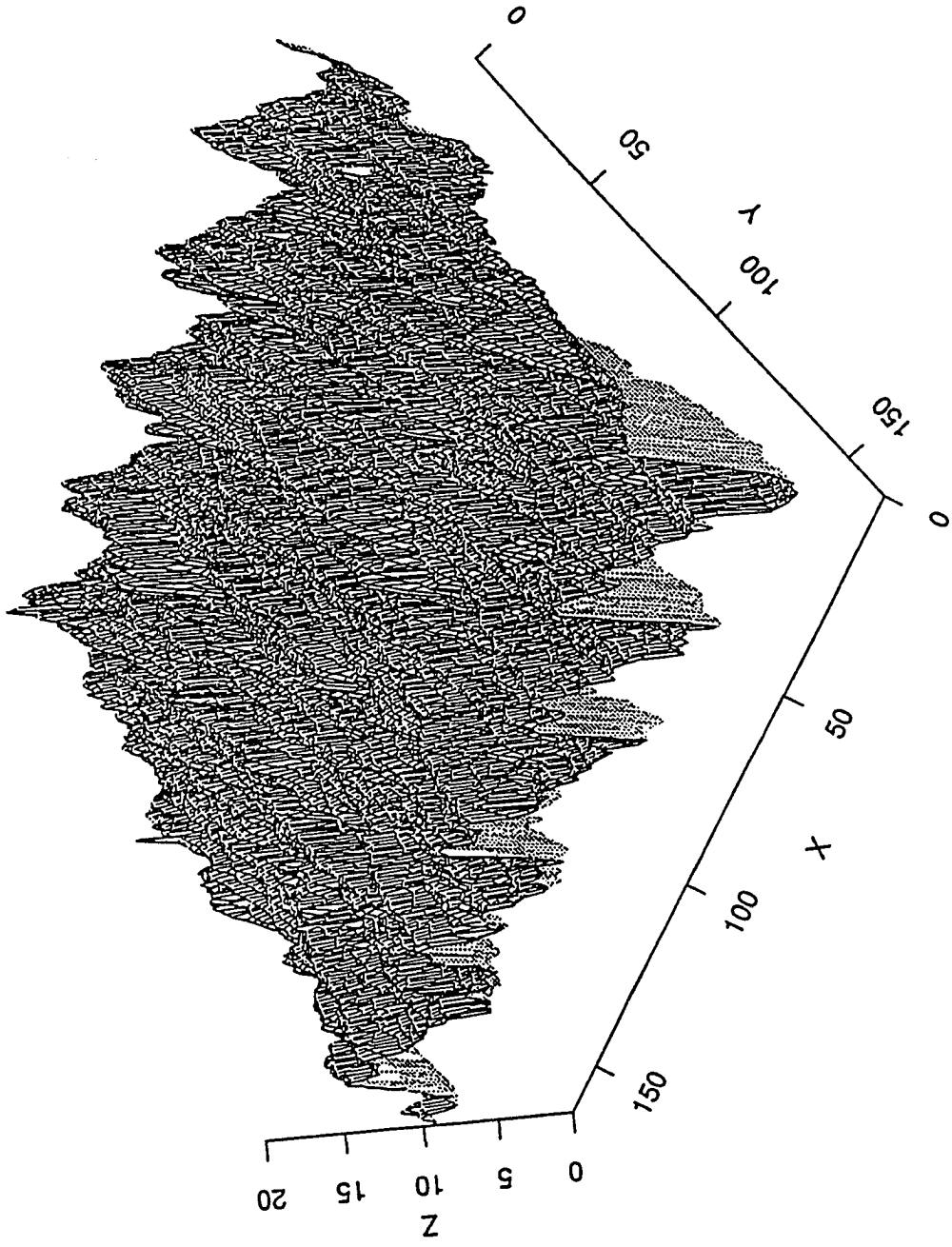


Figure C.2: End milled surface

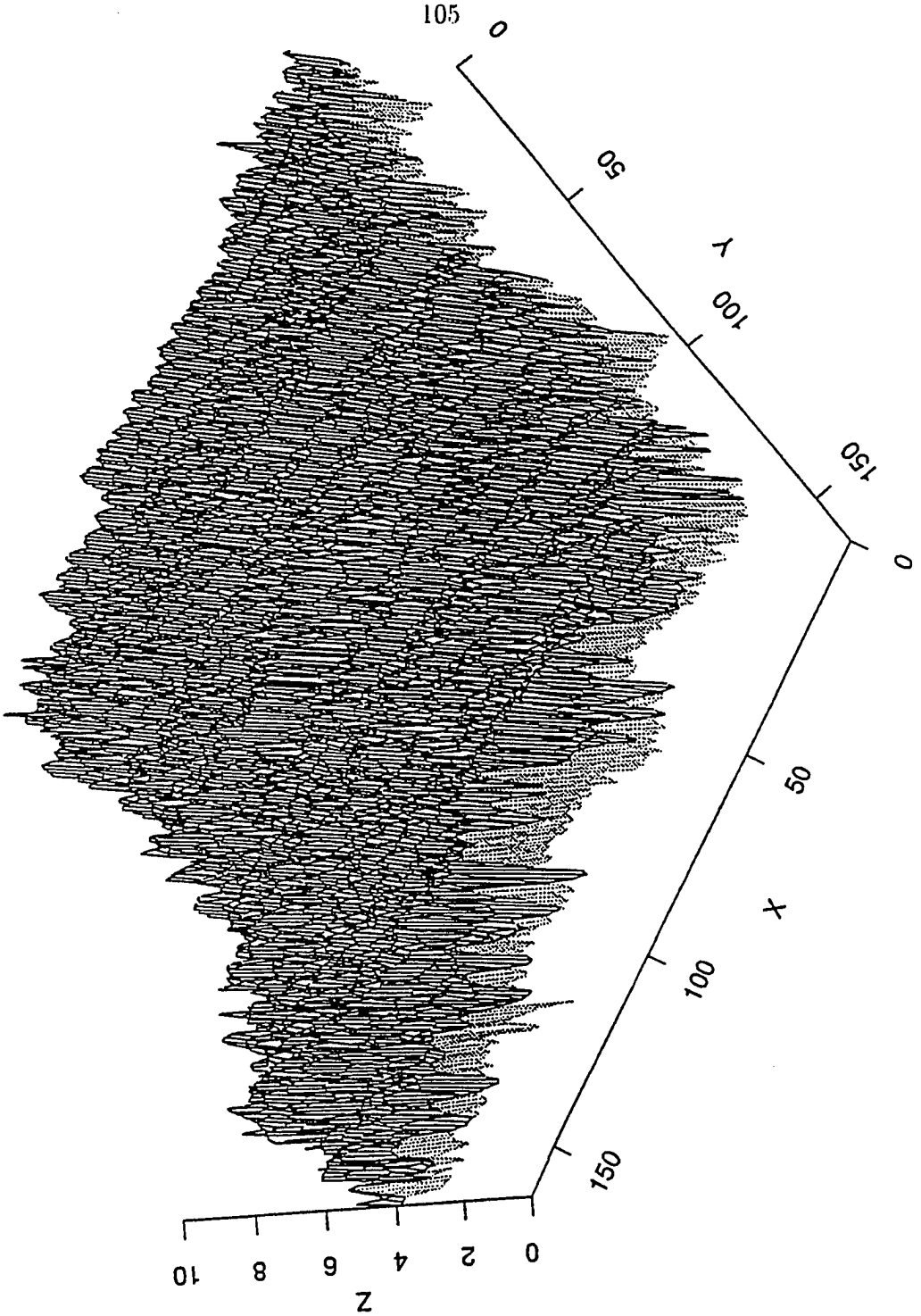


Figure C.3: Fly cut surface

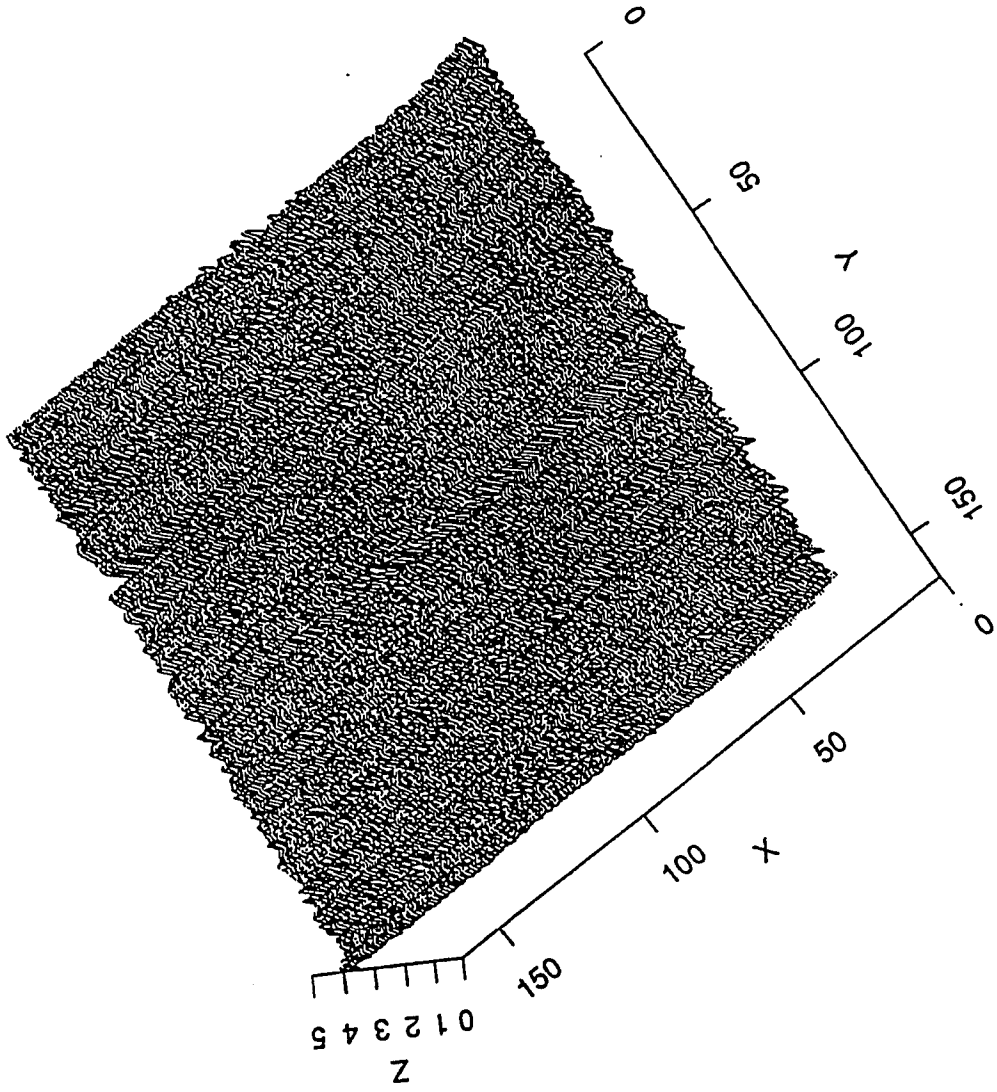


Figure C.4: Ground surface

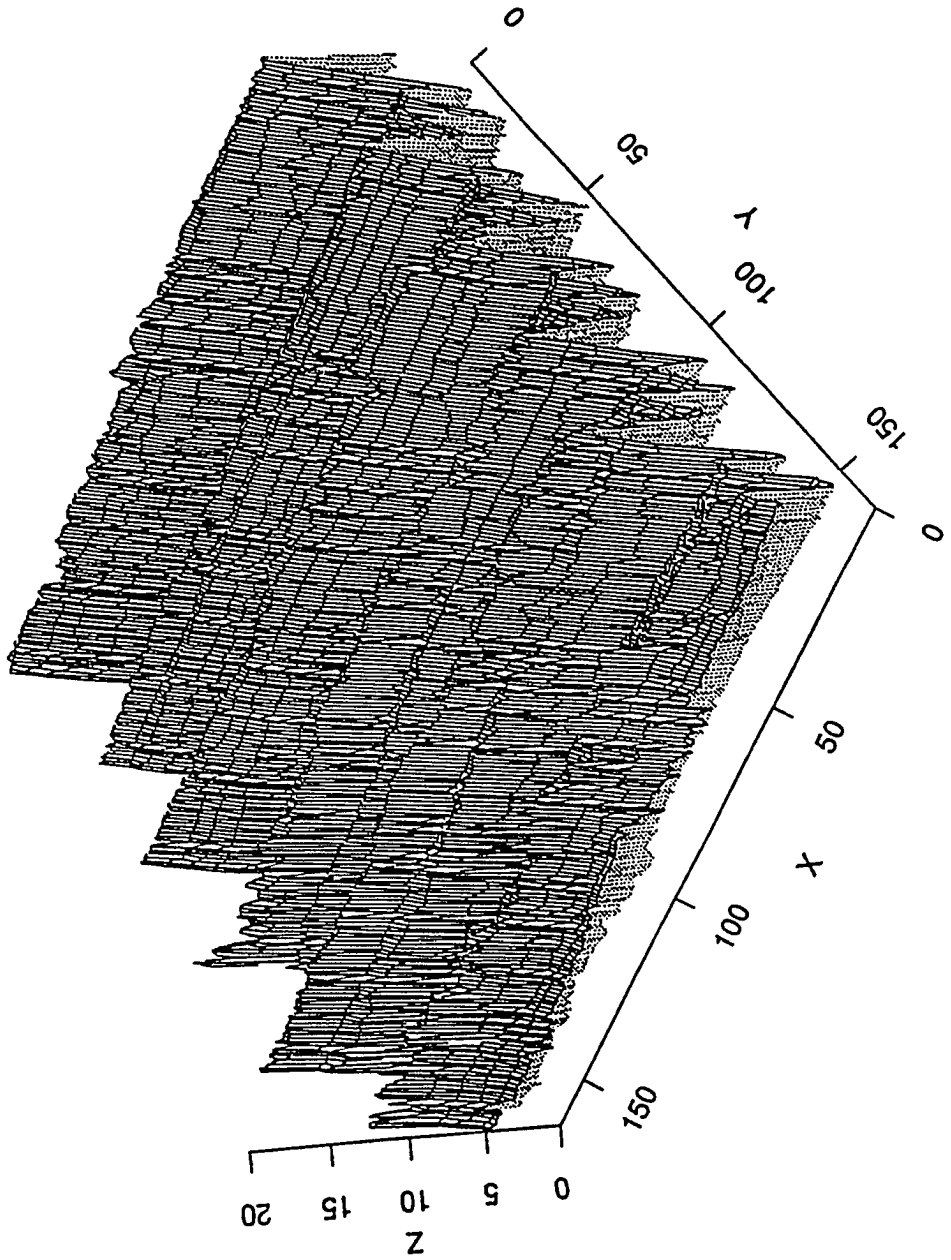


Figure C.5: Shaped surface

APPENDIX D. SPECTRAL PLOTS

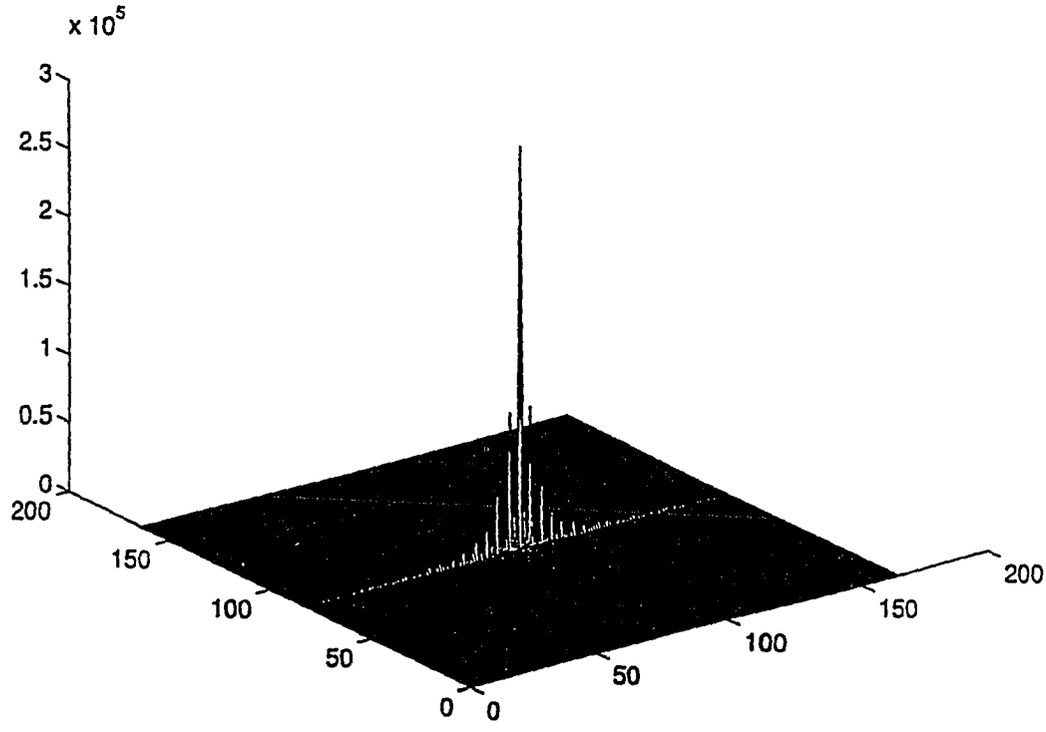


Figure D.1: Spectral plot for bored surface

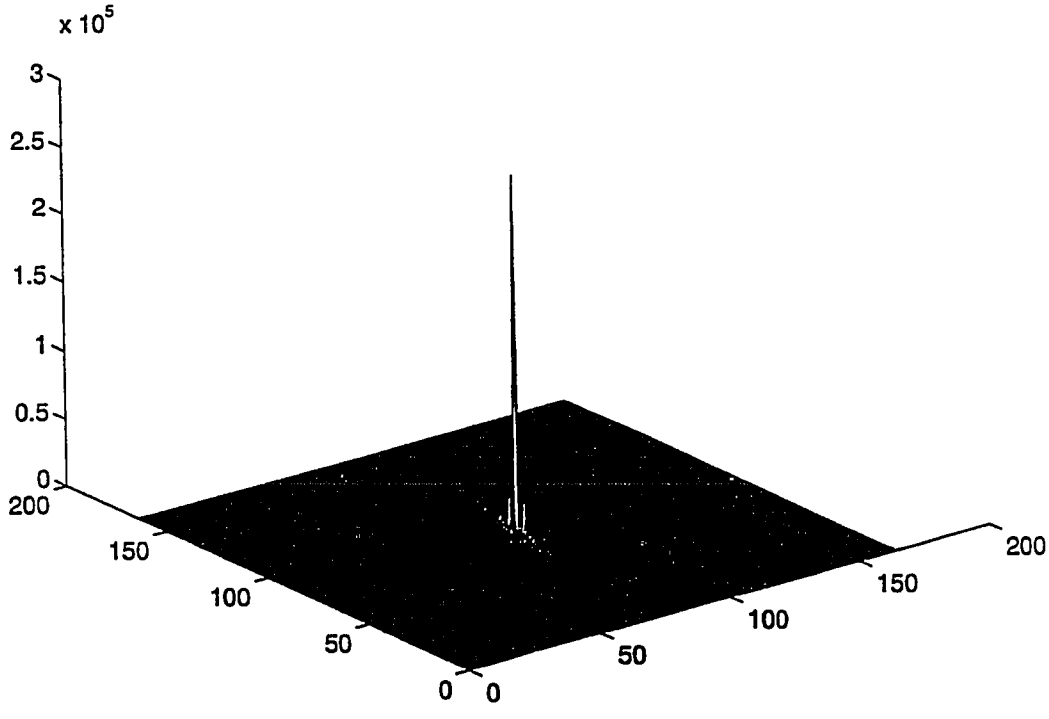


Figure D.2: Spectral plot for end milled surface

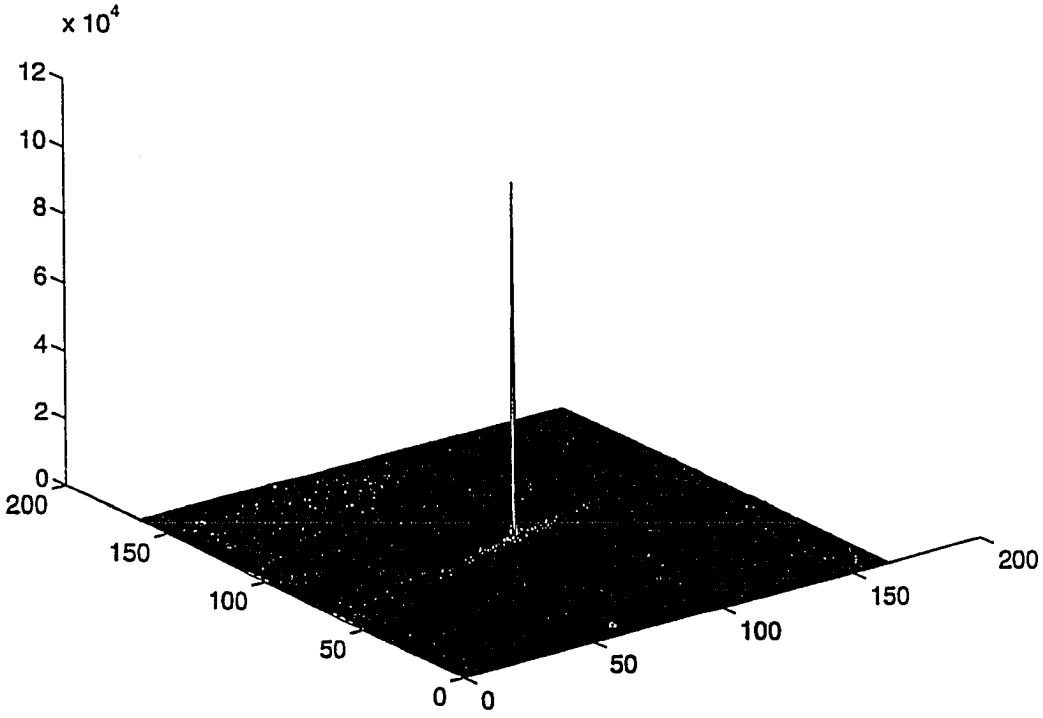


Figure D.3: Spectral plot for fly cut surface

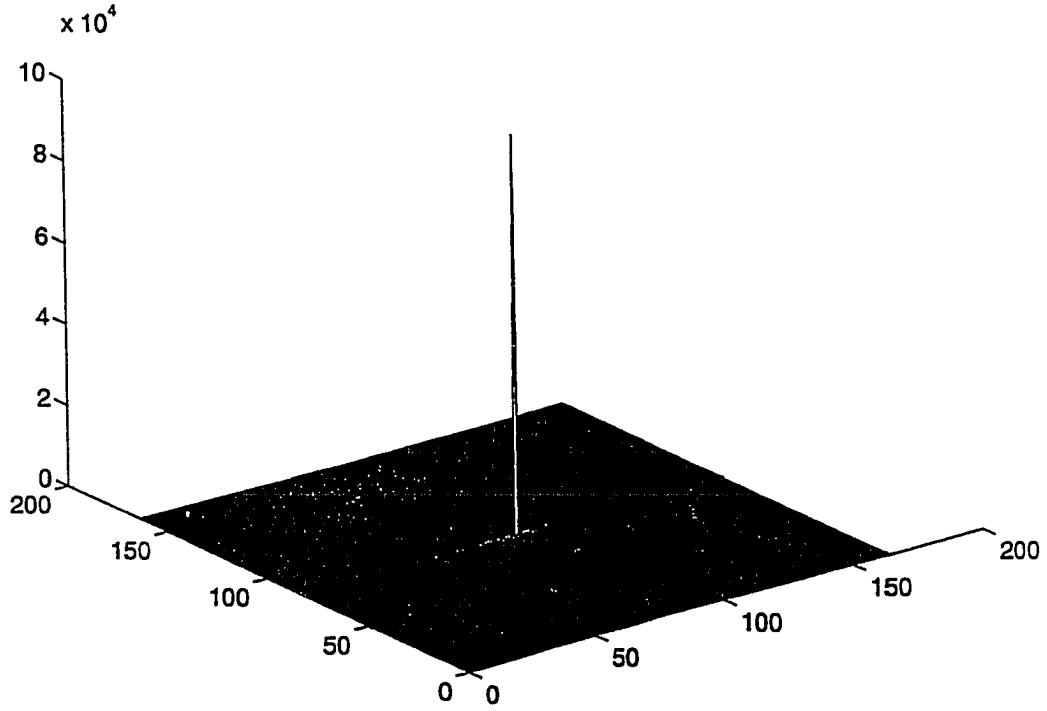


Figure D.4: Spectral plot for ground surface

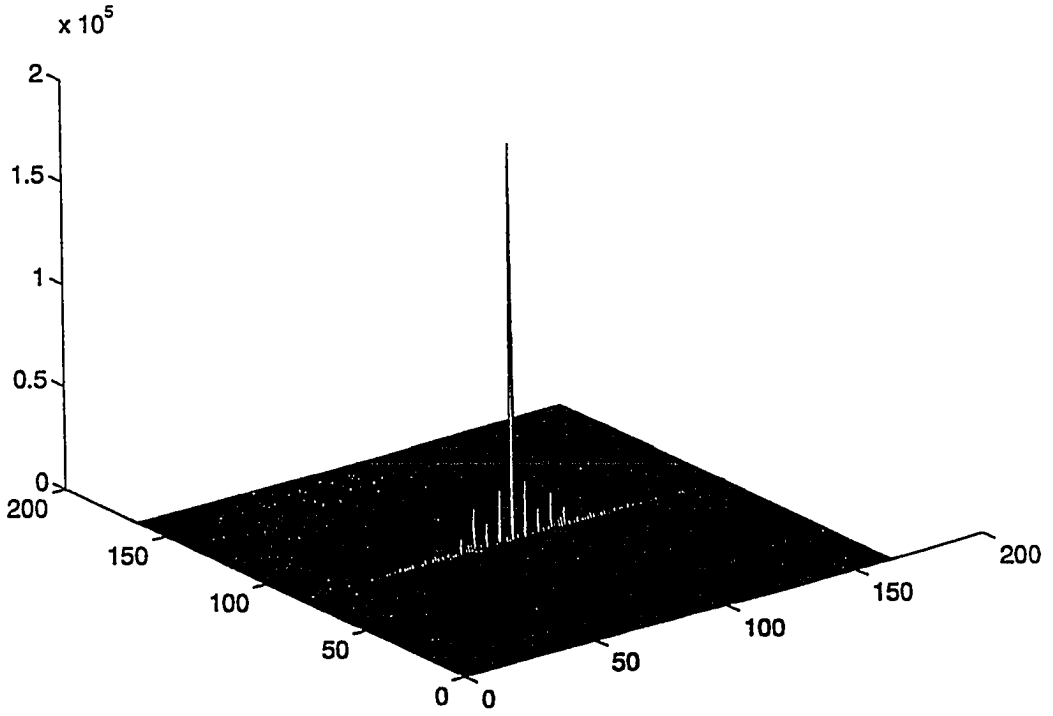


Figure D.5: Spectral plot for shaped surface