

The contact behavior of elastic/plastic non-Gaussian rough surfaces

T.W. Kim^a, B. Bhushan^{a,*} and Y.J. Cho^b

^a*Nanotribology Laboratory for Information Storage and MEMS/NEMS (NLIM), The Ohio State University, 650 Ackerman Road, Suite 255, Columbus, OH 43202-1107, USA*

^b*School of Mechanical Engineering, Pusan National University, San 30 Changjeon-Dong, 609-735, Kumjeong-Gu, Pusan, Korea*

Received 18 July 2005; accepted 21 October 2005; published online 24 May 2006

A three-dimensional contact analysis was conducted to investigate the contact behavior of elastic–perfectly plastic solids with non-Gaussian rough surfaces. The effect of skewness, kurtosis and hardness on contact statistics and the effect of skewness and kurtosis on subsurface stress are studied. Non-Gaussian rough surfaces are generated by the computer with skewness, Sk , of -0.3 , 0.0 and 0.3 , and kurtosis, K , of 2.0 , 3.0 and 4.0 . Contact pressures and subsurface stresses are obtained by contact analysis of a semi-infinite solid based on the use of influence functions and patch solutions. Variation of fractional elastic/plastic contact area, maximum contact pressure and interplanar separation as a function of applied load were studied at different values of skewness and kurtosis. Contact pressure profiles, von Mises stresses, tensile and shear stress contours as a function of friction coefficient were also calculated for surfaces with different skewness and kurtosis. In this study, it is observed that surfaces with $Sk = 0.3$ and $K = 4$ in the six surfaces considered have a minimum contact area and maximum interplanar separation, which may provide low friction and stiction. The critical material hardness is defined as the hardness at which severe level of plastic asperity deformation corresponding to the Greenwood and Williamson's cut-off $A_{\text{plastic}}/A_{\text{real}} = 0.02$ occurs for a given surface and load condition. The critical material hardness of surfaces with $Sk = 0.3$ and $K = 4$ is higher than that of other surfaces considered.

KEY WORDS: contact analysis, kurtosis, surface roughness, skewness

1. Introduction

Surface roughness plays a significant role in friction, wear, and lubrication in machine components. A small change in the distribution of heights, width, and curvature of the asperity peaks can have a noticeable effect on the contact behavior of rough surfaces. In the past four decades, many analytical and numerical models have been developed for contact simulation of rough surfaces [1–9]. The classical analysis is based on the statistical behavior of rough surface contacts with the assumptions of given asperity shapes, height distribution and neglecting interaction between asperities. Greenwood and Williamson [1] analyzed rough surface contacts with the assumption of spherical asperity tips and a Gaussian distribution of asperity heights. Whitehouse and Archard [2] analyzed properties of random surfaces and showed that all significant features of a random surface topography can be defined by two parameters; the rms roughness and the correlation length, if the surface has a Gaussian height distribution. Onions and Archard [3] showed the mean pressure to be nearly independent of load and nominal contact area using Whitehouse and Archard's results. Majumdar and Bhushan [10] and Bhushan and Majumdar [11] developed a new fractal theory of elastic and plastic contact

between two rough surfaces, which uses fractal parameters for surface characterization. Although these statistical models can predict important trends in the effect of surface properties on the real area of contact, their usefulness is limited because of the over-simplified assumptions of asperity geometry, the difficulty in the determination of statistical roughness parameters, and the neglect of interactions between adjacent asperities [6].

With the rapid advance of faster computers within the last decades, more realistic models for contact simulation were developed. Lai and Cheng [12] developed a numerical method to calculate the real area of contact using Patir's [13] computer generated surfaces. Lee and Cheng [14] formulated useful mathematical relationships between the real area of contact, the average gap, and the mean asperity pressure for surfaces with two-dimensional roughness. Tian and Bhushan [5] developed a numerical model that uses digitized real rough surface maps without any arbitrary assumption about the shape or size of asperities. They utilized the approach on the basis of the variational principle. Their results were based on a numerical contact simulation model which has an advantage over the classical approach in that the numerical models closely duplicate the actual surface contact by taking into account the elastic interactions between the asperities.

In most models, a Gaussian distribution is assumed. However, the assumption that most surface heights have

*To whom correspondence should be addressed.
E-mail: Bhushan.2@osu.edu

a Gaussian distribution is not strictly true. Most of the common machining processes produce surfaces with non-Gaussian distribution [7, 15–17]. Turning, shaping and electrodischarge machining produce a peaked surface with positive skewness. Grinding, honing, milling and abrasion processes produce grooved surfaces with negative skewness but high kurtosis values. Laser polishing produces surfaces with high kurtosis. Disk surfaces produced by sputter texturing [18] have mainly peaks on the disk surfaces which are expected to have positive skewness.

Kotwal and Bhushan [19] developed an analytical model for contact analyses of non-Gaussian surfaces. Chilamakuri and Bhushanm [20] have developed a method to produce non-Gaussian rough surfaces on the computer and performed numerical contact analyses on the surfaces with various skewness and kurtosis. They found that a surface with slight positive skewness and kurtosis greater than four results in an optimum surface with a minimum contact area and meniscus force. However, they did not analyze the subsurface stress fields and the effect of material hardness.

In the present study, a three-dimensional contact analysis was conducted to investigate the contact behavior of elastic–perfectly plastic solids with non-Gaussian distribution of surface heights. Three-dimensional non-Gaussian rough surfaces were generated on the computer with different values of skewness and kurtosis. Variation of fractional contact area, plastic area, maximum contact pressure and interplanar separation as a function of applied load and material hardness were studied at different values of skewness and kurtosis. Contact pressure profiles, von Mises stresses, tensile and shear stress contours were also calculated for non-Gaussian surfaces in both frictionless and frictional contacts.

2. Numerical Generation of Non-Gaussian Random Surface

In most previous studies, the surface parameters expressed as the root mean square (RMS) roughness (σ) or the center line average roughness (Ra) are used. This method neglects other surface descriptions. Though surface profiles may have the same σ or Ra, they could have different asperity distributions. Thus, skewness (Sk) and kurtosis (K) which give information about the shape of asperities distribution should be used to represent a more exact surface profile. The probability density function $p(z)$ for surface heights gives the probability of locating a point at a height z . This shape can be described in terms of moments of the function about the mean, referred to as central moments. The first moment is m , which is taken to be zero and the second moment is the variance σ^2 , the square of the standard deviation,

$$\sigma^2 = \int_{-\infty}^{\infty} z^2 p(z) dz \quad (1)$$

The third moment is the skewness, which represents asymmetric spread of the height distribution and is given by

$$Sk = \frac{1}{\sigma^3} \int_{-\infty}^{\infty} z^3 p(z) dz \quad (2)$$

where $p(z)$ is the probability density function of surface heights, z . A Gaussian surface has zero skewness with an equal number of peaks and valleys at certain heights. Profiles with peaks removed have negative skewness whereas profiles with valleys filled or high peaks have positive skewness. The fourth moment is termed kurtosis, which represents the peakedness of the distribution and is given by

$$K = \frac{1}{\sigma^4} \int_{-\infty}^{\infty} z^4 p(z) dz \quad (3)$$

The kurtosis shows the degree of pointedness or bluntness of waveform. The kurtosis always has a positive value and measures the sharpness of symmetric distribution. For a Gaussian distribution the curve has a kurtosis of 3. If $K < 3$, the distribution has relatively few high peaks and low valleys. If $K > 3$, the distribution has more high peaks and low valleys.

Non-Gaussian surfaces with various Sk and K values are generated on the computer using a two-dimensional digital filter technique [21]. A digital filter technique is a system that transforms an input sequence $\eta(I, J)$ into an output sequence $z(I, J)$, the digitized surface heights. First an input sequence of independent random numbers having a Gaussian distribution of a known standard deviation, $\eta(I, J)$, is generated using a random number generator. To simulate real rough surfaces, however, surfaces having an expected autocorrelation function (ACF) and height distribution need to be generated. The two-dimensional digital filter technique is employed to generate an output sequence $z(I, J)$ of a known autocorrelation by a linear transformation system defined as

$$z(I, J) = \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} h(k, l) \eta(I - k, J - l) \quad (4)$$

where $I = 0, 1, \dots, N-1$, $J = 0, 1, \dots, M-1$, $n = N/2$, $m = M/2$ and h is a filter function. The Fourier transform of equation (4) is given by

$$Z(\omega_x, \omega_y) = H(\omega_x, \omega_y) A(\omega_x, \omega_y) \quad (5)$$

where A and Z are Fourier transformations of the input η and output sequences z respectively, and H is the transfer function of the system given by

$$H(\omega_x, \omega_y) = \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} h(k, l) e^{-jk\omega_x - jl\omega_y} \quad (6)$$

The ACF of a random surface in the discrete form is defined as

$$R_z(k, l) = \frac{1}{NM} \sum_{I=0}^{N-1} \sum_{J=0}^{M-1} z(I, J)z(I+k, J+l) \quad (7)$$

The power spectral density (PSD) function of the surface is the Fourier transform of R_z , given by

$$S_z(\omega_x, \omega_y) = \frac{1}{nm} \sum_{k=-n/2+1}^{n/2-1} \sum_{l=-m/2+1}^{m/2-1} R_z(k, l)e^{-jk\omega_x - jl\omega_y} \quad (8)$$

If $S_\eta(\omega_x, \omega_y)$ denotes the PSD function of the input sequence, the relation between S_η and S_z for a linear system has the form

$$S_z(\omega_x, \omega_y) = |H(\omega_x, \omega_y)|^2 S_\eta(\omega_x, \omega_y) \quad (9)$$

Since η is a sequence composed of independent random number, its PSD function is a constant. The inverse Fourier transform of $H(\omega_x, \omega_y)$ obtained from equation (9) gives the filter function h :

$$h(k, l) = \frac{1}{nm} \sum_{\omega_x=-n/2+1}^{n/2-1} \sum_{\omega_y=-m/2+1}^{m/2-1} H(\omega_x, \omega_y)e^{-jk\omega_x - jl\omega_y} \quad (10)$$

Now $h(k, l)$ from equation (10) is used in equation (4) to obtain $z(I, J)$.

So far, a technique has been described to generate Gaussian random surfaces. For the generation of non-Gaussian random surfaces, the Gaussian input sequence is transformed to an input sequence with appropriate skewness, Sk_η , and kurtosis, K_η , by using the Johnson translator system of distribution [22,23]. However, during later transformation using the two-dimensional filter technique to obtain an output sequence, the distribution of the output sequence in most cases will not be of the same form as that of the input sequence. In this case, the skewness and kurtosis of the output sequence are related to those of the input sequence as

$$(Sk_z)^{1/2} = \frac{\sum_{i=0}^q \theta_i^3}{(\sum_{i=0}^q \theta_i^2)^{3/2}} (Sk_\eta)^{1/2} \quad (11)$$

$$(K_z)^{1/2} = \frac{\sum_{i=0}^q \theta_i^4 K_\eta + \sum_{i=0}^{q-1} \sum_{j=i+1}^q \theta_i^2 \theta_j^2}{(\sum_{i=0}^q \theta_i^2)^2} \quad (12)$$

where Sk_z and K_z are the required skewness and kurtosis, Sk_η and K_η are the input skewness and kurtosis for Johnson's translator system, and

$$\begin{aligned} \theta_i &= h(k, l), \quad k = 0, Y, n-1 \\ l &= 0, Y, m-1 \quad i = km+1 \end{aligned} \quad (13)$$

The non-Gaussian input sequence with modified skewness and kurtosis obtained from equations (11) and (12) is generated using Johnson's translator system. The

Johnson system of frequency curves based on the method of moments provides a family of curves that can be used to generate an equation for the distribution for which the first four moments (mean, σ , Sk and K) are known. In this system there are three main types of frequency curves, S_U , S_L and S_B , defined as

$$S_U: \quad \eta' = \gamma + \delta \sinh^{-1} \left(\frac{\eta'' - \xi}{\lambda} \right) \quad (14)$$

$$S_L: \quad \eta' = \gamma + \delta \log \left(\frac{\eta'' - \xi}{\lambda} \right) \quad (\eta'' > \xi) \quad (15)$$

$$S_B: \quad \eta' = \gamma + \delta \log \left(\frac{\eta'' - \xi}{\xi + \lambda - \eta''} \right) \quad (\xi < \eta'' < \xi + \lambda) \quad (16)$$

where η' is a standardized normal variable sequence, η'' is the variable sequence with given skewness and kurtosis, and γ , δ , ξ and λ are constants to be determined for the first four given moments by using method of moments. Finally, steps of the two-dimensional digital filter technique outlined earlier by equations (6)–(12) are used to transform the non-Gaussian input sequence η'' to z .

To summarize, non-Gaussian surfaces are generated using the following steps:

1. Generate a Gaussian input sequence $\eta'(I, J)$ using a random number generator.
2. Determine the skewness and kurtosis of the input sequence for desired skewness and kurtosis of an output sequence, using equation (11) and (12).
3. Transform the Gaussian input sequence to the input sequence $\eta''(I, J)$, which has a distribution with appropriate skewness, Sk_η , and kurtosis, K_η , by using the Johnson translator system of distribution.
4. Assume that the ACF for the simulated surface has the form

$$R_z(k, l) = \sigma^2 \exp \left[-2.3 \sqrt{\left(\frac{k}{\beta_x^*} \right)^2 + \left(\frac{l}{\beta_y^*} \right)^2} \right] \quad (17)$$

where σ is the standard deviation of surface heights and β_x^* , β_y^* are the correlation lengths of surface profiles in the x and y directions. An isotropic surface is considered, i.e., $\beta_x^* = \beta_y^* = \beta^*$.

5. Calculate $S_z(\omega_x, \omega_y)$ using equations (17) and (8).
6. For the known values of S_η and S_z , calculate $H(\omega_x, \omega_y)$ using equation (9)
7. Obtain $h(k, l)$ by the inverse Fourier transform of $H(\omega_x, \omega_y)$ using equation (10)
8. Obtain $z(I, J)$ by performing the filtering operation using equation (4)

Figure 1 shows two- and three-dimensional surfaces generated according to the values of kurtosis and

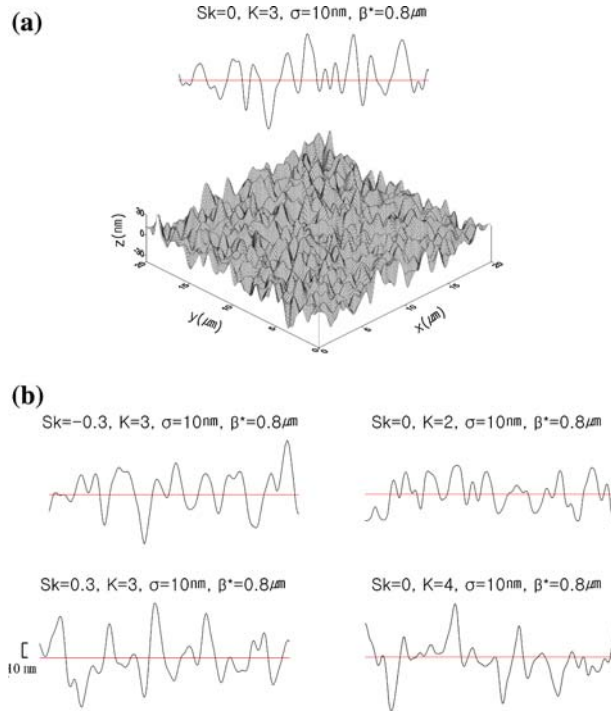


Figure 1. Surface height maps of computer-generated rough surface, (a) three-dimensional surface maps with Gaussian distribution and (b) two-dimensional surface profiles of non-Gaussian surfaces with different skewness and kurtosis.

skewness using this method. For the contact analysis, surfaces having same correlation lengths $\beta^* = 0.8\mu\text{m}$ and $\sigma = 10$ with different values of skewness and kurtosis were generated using a scan size (profile length in the x and y directions) of $20 \times 20 \mu\text{m}$. Correlation length is defined as the distance at which the value of the ACF reduces to 10% of the value at the origin.

3. Contact Analysis

The rough surface is modeled as a semi-infinite space with computer-generated surface topography. The computer-generated surface contains 256×256 data points. By using these data points as nodal points there will be 255×255 small patches of equal size ($0.078 \times 0.078 \mu\text{m}^2$) within the simulated area. This data file has the similar data format of a measured surface map from an atomic force microscope or a non-contact optical profiler.

In this study, matrix inversion technique and iterative process are used to obtain contact pressure distributions, real contact areas and interplanar separations [24]. This technique takes full account of the interaction of deformation from all contact points and predicts the contact geometry of real surfaces under loading. It provides useful information on the contact pressure, number of contacts, their sizes and distributions, and interplanar separations. The matrix

inversion approach starts with the classical approach for finding the relationship between the contact pressure $p(x,y)$ and surface displacement $u(x,y)$ by Boussinesq [25] who made use of the theory of potential. Generally, the contact of two elastic bodies can be simulated as the contact of a rigid body pressing down on an elastic half space. Under the applied load the total gap $f(x,y)$ after deformation is

$$f(x,y) = e(x,y) + u(x,y) - d \quad (18)$$

where $e(x,y)$ is the shape function before contact, $u(x,y)$ is the elastic deformation, and d is the effective rigid plane displacement. The condition of $f(x,y) = 0$ in the contact domain Ω must be satisfied, therefore equation (18) leads to the following equation in terms of Boussinesq's solution.

$$\frac{1 - \nu^2}{\pi E} \iint_{\Omega} \frac{p(x',y') dx' dy'}{\sqrt{(x-x')^2 + (y-y')^2}} = d - e(x,y) \quad (19)$$

where E is Young's modulus and ν is Poisson's ratio. This equation can be written in the following form.

$$\sum_{k=1}^M C_{kl} p_k = d - e_k \quad (20)$$

where M is the number of patches in Ω , and C_{kl} is the influence matrix, which represents the displacement at point k due to a distributed unit normal load on element l . The numerical procedure for finding the final contact solution starts with a given d value and an initial guess of nodes assumed to be in contact. The pressures are solved by inverting equation (20). After calculating the pressures, the deformations are also calculated using the same equation. The pressures and deformations are continually iterated until convergence occurs to all the surface elements with positive or zero contact pressure.

The elastic contact model is extended to elastic-plastic contact analysis for the case where the solid in contact deforms as an elastic-perfectly plastic material. It is assumed that the region of the plastic deformation is confined to a very small area and does not significantly alter the geometry of the elastically deformed contact surface. Outside the plastic region, the relationship between contact pressure and elastic deformation still holds. Since the contacting asperities at moderate loads are normally separated from each other in rough flat surface contact and are composed of few patches, these are much like sharp asperities instead of a spherical asperity of large radius. These points are assumed to be elastic-perfectly plastic, which transfers from elastic to fully developed plastic flow immediately. Therefore, it is a constrained plastic deformation rather than fully developed plastic flow situation. This approach will generate errors in analyzing the contact of a sphere and heavily loaded contact of rough flat surface.

In modeling the plastic deformation behavior of the asperities, if any local contact region experiences pressures exceeding three times the uniaxial yield strength, $3Y$ or hardness of the softer of the two mating bodies, then the contact pressures in those regions were set equal to $3Y$ and the regions were allowed to deform freely [14,26]. Thus, in this study, once the contact pressure exceeds the hardness of the softer material, H ($\sim 3Y$), the stress is set to be equal to the hardness.

Once the contact pressures and locations of the contact points are known, the sub-surface stress field can be calculated. A closed-form solution of the interior stress field has been derived for the case of a uniform normal stress distribution, p_k , over a rectangular patch of length 2α in x direction and 2β in y direction on the surface of a semi-infinite elastic space [27] (figure 2), given by

$$\sigma_{ij}^k = \frac{p_k}{2\pi} [A_{ij}(\bar{x} + \alpha, \bar{y} + \beta, \bar{z}) + A_{ij}(\bar{x} - \alpha, \bar{y} - \beta, \bar{z}) - A_{ij}(\bar{x} - \alpha, \bar{y} + \beta, \bar{z}) - A_{ij}(\bar{x} + \alpha, \bar{y} - \beta, \bar{z})] \quad (21)$$

where $\bar{x} = x - x_k$, $\bar{y} = y - y_k$, $\bar{z} = z$ and $(x_k, y_k, 0)$ is the coordinate of the center of k th patch. The functions A_{ij} are given as

$$\begin{aligned} A_{xx}(x, y, z) &= 2v \left[\tan^{-1} \left(\frac{xz}{yR} \right) - \tan^{-1} \left(\frac{x}{y} \right) \right] \\ &\quad - \tan^{-1} \left(\frac{y}{x} \right) + \tan^{-1} \left(\frac{yz}{xR} \right) + \frac{xyz}{R(x^2 + z^2)} \\ A_{yy}(x, y, z) &= 2v \left[\tan^{-1} \left(\frac{yz}{xR} \right) - \tan^{-1} \left(\frac{y}{x} \right) \right] - \tan^{-1} \left(\frac{x}{y} \right) \\ &\quad + \tan^{-1} \left(\frac{xz}{yR} \right) + \frac{xyz}{R(y^2 + z^2)} \\ A_{zz}(x, y, z) &= -\tan^{-1} \left(\frac{y}{x} \right) - \tan^{-1} \left(\frac{x}{y} \right) + \tan^{-1} \left(\frac{yz}{xR} \right) \\ &\quad + \tan^{-1} \left(\frac{xz}{yR} \right) - \frac{xyz}{R(x^2 + z^2)} - \frac{xyz}{R(y^2 + z^2)} \\ A_{xy}(x, y, z) &= -(1 - 2v) \ln(R + z) - \frac{z}{R} \\ A_{xz}(x, y, z) &= \frac{z^2 y}{R(x^2 + z^2)} \\ A_{yz}(x, y, z) &= \frac{z^2 x}{R(y^2 + z^2)} \end{aligned} \quad (22)$$

where $R = \sqrt{x^2 + y^2 + z^2}$.

Contributions from all patch loading are superimposed, so the stresses are

$$\sigma_{ij} = \sum_{k=1}^M \sigma_{ij}^k \quad (23)$$

The previously obtained contact points and pressures are used as input; since this information is obtained at a mesh of 255×255 patches with an equal size of $0.078 \times 0.078 \mu\text{m}^2$, the same patch size is used in the patch solution. Note that the patch solution is for elastic

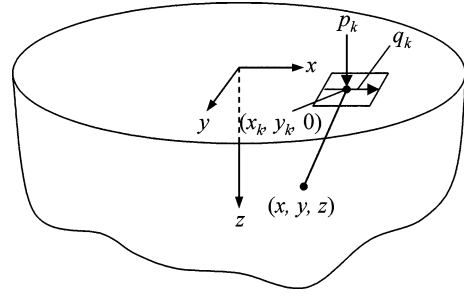


Figure 2. Illustration of a semi-infinite space subject a rectangular patch loading on the surface with uniform normal and tangential loads.

analysis. When plastic deformation is initiated, this method cannot be applied to calculating the sub-surface stress. Nevertheless, the elastic solution can be used to determine if yielding is reached in the bulk material (especially when the contact pressure is equal to the hardness, $p = H$) and to provide basic information for elastic-plastic stress analysis.

The effect of friction on the contact pressures and contact locations has been neglected, so the tangential load at each contact point is equal to its contact load multiplied by a coefficient of friction, μ . The interior stress field induced by a uniformly distributed tangential load, q_k , over a rectangular patch of length 2α in x direction and 2β in y direction on the surface of an elastic semi-infinite space has also been derived [28] (figure 2) by

$$\sigma_{ij}^k = \frac{q_k}{2\pi} [B_{ij}(\bar{x} + \alpha, \bar{y} + \beta, \bar{z}) + B_{ij}(\bar{x} - \alpha, \bar{y} - \beta, \bar{z}) - B_{ij}(\bar{x} - \alpha, \bar{y} + \beta, \bar{z}) - B_{ij}(\bar{x} + \alpha, \bar{y} - \beta, \bar{z})] \quad (24)$$

where q_k is the uniform tangential loading applied over the k th patch. The functions B_{ij} are given as

$$\begin{aligned} B_{xx}(x, y, z) &= 2 \ln(R + y) + \frac{x^2 y}{R(x^2 + z^2)} + (1 - 2v) \frac{y}{R + z} \\ B_{yy}(x, y, z) &= 2 \ln(R + y) - \frac{y}{R} + (1 - 2v) \frac{y}{R + z} \\ B_{zz}(x, y, z) &= \frac{z^2 y}{R(x^2 + z^2)} \\ B_{xy}(x, y, z) &= \ln(R + x) - \frac{x}{R} + (1 - 2v) \frac{x}{R + z} \\ B_{xz}(x, y, z) &= -\tan^{-1} \left(\frac{xy}{Rz} \right) + \frac{xyz}{R(x^2 + z^2)} \\ B_{yz}(x, y, z) &= -\frac{z}{R} \end{aligned} \quad (25)$$

Contributions from all patch loadings are superimposed. The complete subsurface stress field for frictional contact is obtained from the superposition of the normal and tangential loading solutions. Note that this method does not account for transverse strain or junction growth. Understanding of junction growth starts with the deformation of individual asperities under the

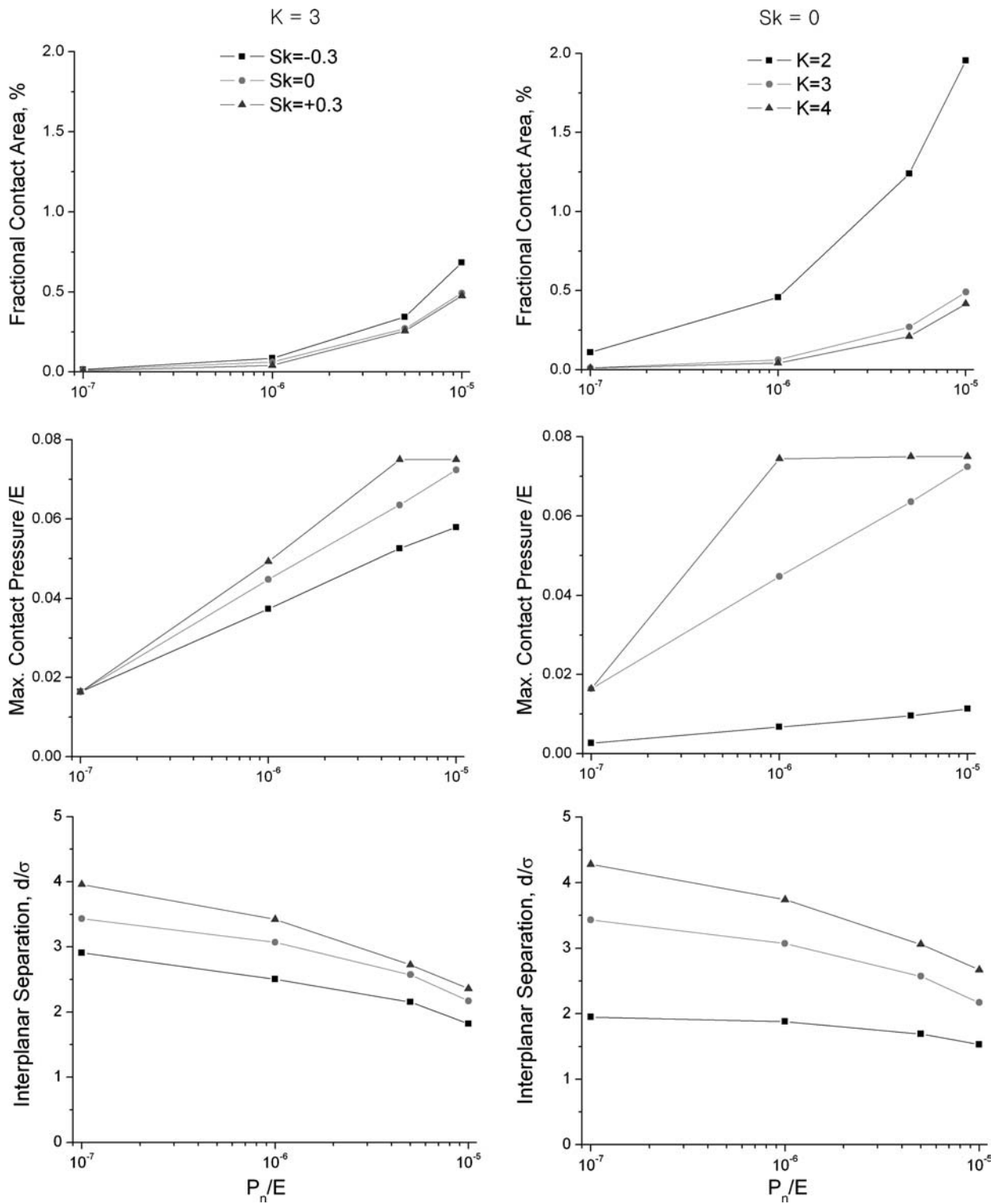


Figure 3. Variation of the fractional contact area, the maximum contact pressure and the interplanar separation for surfaces with different skewness and kurtosis as a function of applied load.

combined action of both normal and shear stresses. If the local shear stress increases then the associated normal pressure necessary to maintain plasticity of the junction reduces; if the load stays the same then a

reduction in pressure demands an increase in the true area supporting the load. Therefore the true contact area can be greater than what this approach would predict, which impairs the accuracy of the results.

4. Results and Discussion

This analysis was carried out for a rigid surface in both frictionless and frictional contacts with a non-Gaussian rough elastic-perfectly plastic solid surface. The upper surface is assumed to be rigid to simplify the problem. Elastic modulus of softer rough surface E is taken as 100 GPa and Poisson's ratio is taken as 0.3. The pressure is normalized by E . Non-Gaussian rough surfaces are generated by the computer with skewness, Sk , of -0.3, 0.0 and 0.3, and kurtosis, K , of 2.0, 3.0 and 4.0. Rough surfaces are analyzed with the applied load P_n/E ranging from 10^{-7} to 10^{-5} and hardness H/E ranging from 0.008 to 0.1. Two values of the coefficient of friction are examined for frictional contact, $\mu = 0.25$ and $\mu = 0.5$.

4.1. Effect of Skewness and Kurtosis

Figure 3 shows the fractional contact area, the maximum contact pressure and the interplanar separation

for surfaces with different skewness and kurtosis as a function of applied load. For the case of Gaussian surface ($Sk = 0.0$ and $K = 3.0$), the trends of the real contact area and the maximum contact pressure change are consistent with that of Peng and Bhushan [29]. As shown in this figure, a surface with $Sk = -0.3$ of the three values of Sk considered gives the largest fractional contact area and the lowest maximum contact pressure. This is not very clear at lower normal load because there are only a few contact points at this load. However, at higher loads, this is clear. Surfaces with $Sk = 0$ and $Sk = 0.3$ give similar contact area results at all loads, but a surface with $Sk = 0.3$ gives a higher maximum contact pressure. Surfaces with $Sk = 0$ and 0.3 consist of the adequate number of peaks to support the applied load, whereas a surface with a negative skewness has a larger number of peaks at a certain height with deep valleys that afford a high contact area and low contact pressure. The interplanar separation means a distance between a central line of rough surface and a rigid flat plate. The interplanar separation has a maximum value

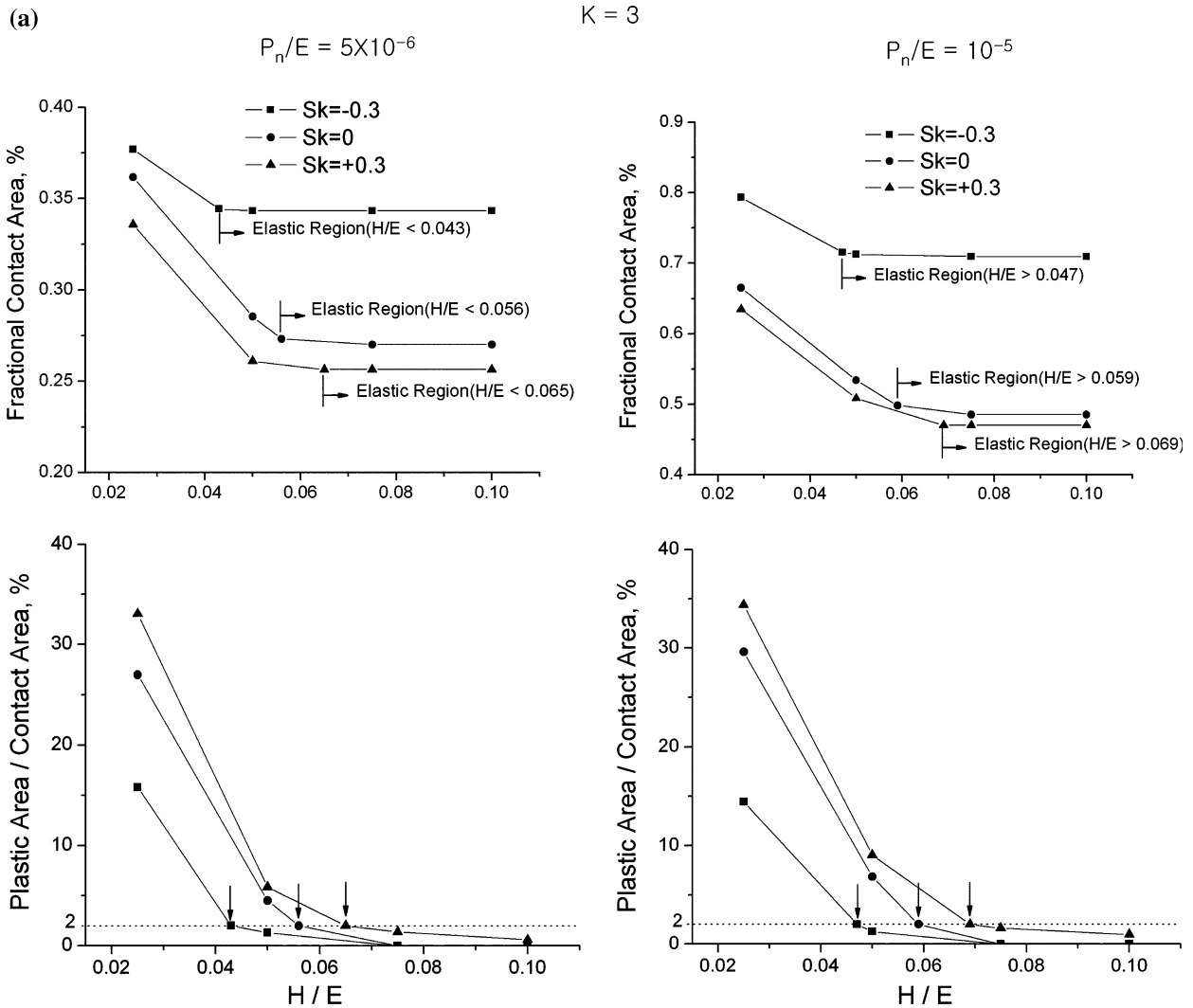


Figure 4. Variation of the fractional contact area and a ratio of plastic area to contact area as a function of material hardness for $P_n/E = 5 \times 10^{-6}$ and 10^{-5} , (a) skewness and (b) kurtosis.

(b)

Sk = 3

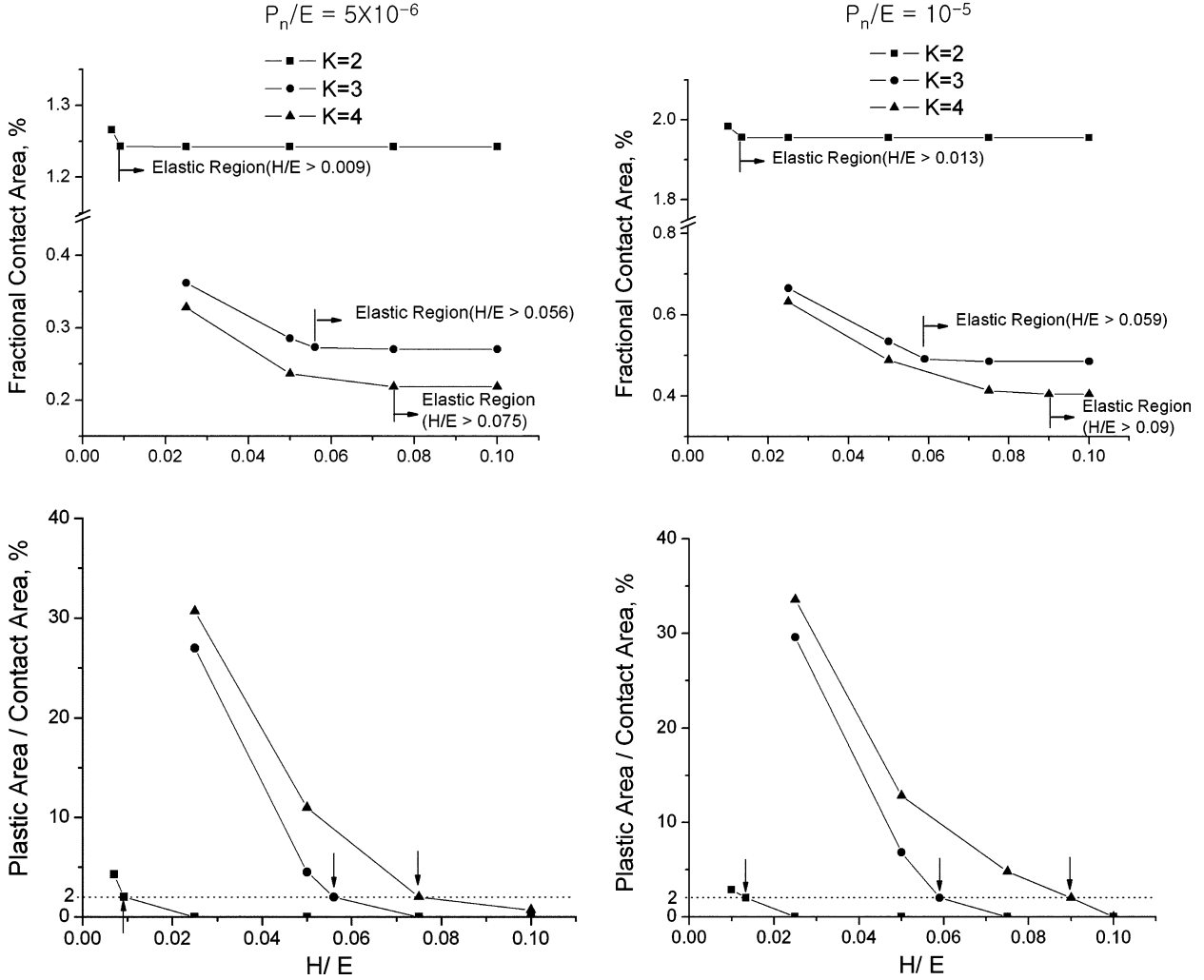


Figure 4. Continued.

at surface with $Sk = 0.3$ among the three values of Sk considered, but the effect of skewness decreases with an increase in the normal load. The fractional contact area and the interplanar separation decrease as the kurtosis value increases. However, the maximum pressure increases with kurtosis due to a decrease in the contact area with kurtosis. An increase in the peak-to-mean distance and a decrease in the peak density with an increase in kurtosis result in a lower contact area and a higher contact pressure.

4.2. Effect of hardness

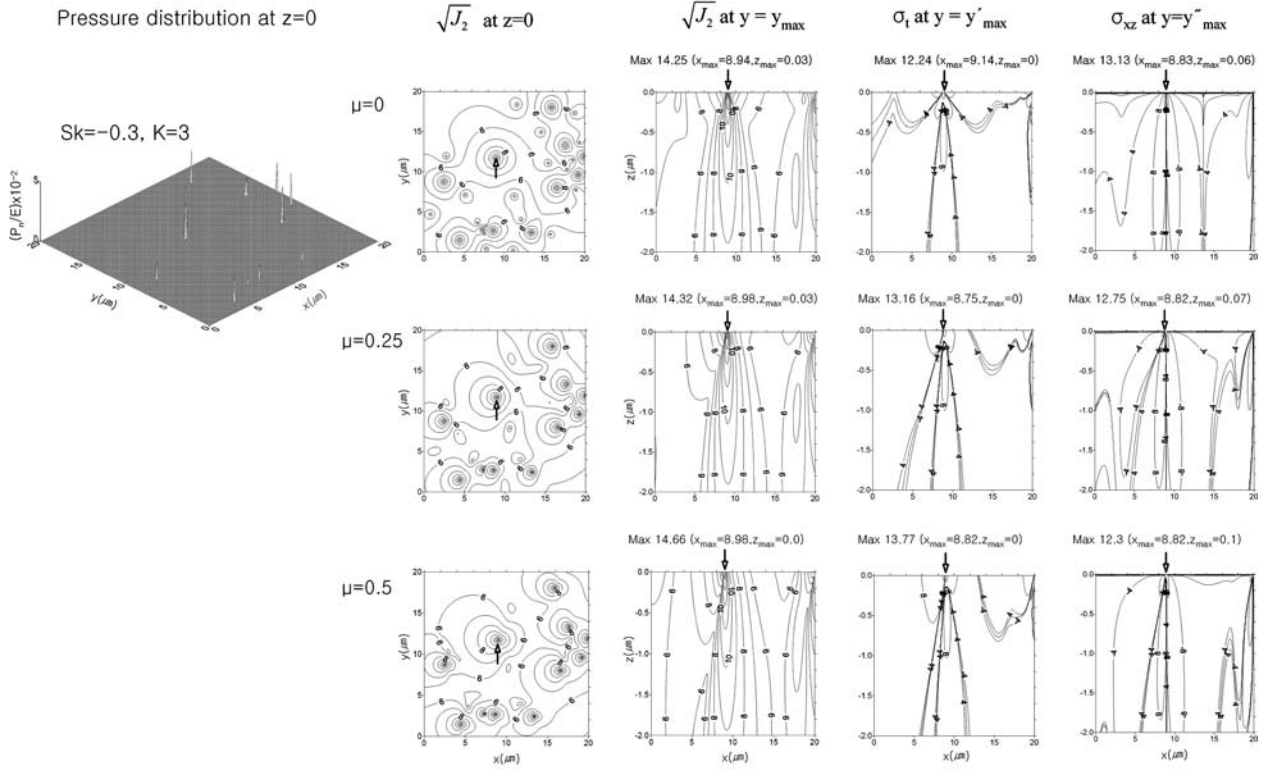
Figure 4 shows the fractional contact area and a ratio of plastic area to real contact area for surfaces with different skewness and kurtosis as a function of material hardness at $P_n/E = 5 \times 10^{-6}$ and 10^{-5} . For the limit of elastic deformation, Greenwood and Williamson [1]

defined a plasticity index ψ . They found that if $\psi < 0.6$, the deformation is largely elastic and if $\psi > 1$, surface deformation is largely plastic, assuming $A_{plastic}/A_{real} = 0.02$ to be the criterion for the onset of a significant degree of plasticity. Therefore, in this study, the critical material hardness is defined as the value of H/E at which severe level of plastic asperity deformation corresponding to the Greenwood and Williamson's cutoff $A_{plastic}/A_{real} = 0.02$ occurs for a given surface and load condition, which indicates points of inflection where fractional contact areas are not reduced any more with an increase in H/E in figure 4. Thus, in figure 4, elastic region means a region satisfying $A_{plastic}/A_{real} < 0.02$.

For all the case of figure 4, it is observed that the real contact area becomes independent of the material hardness and approaches the results in the pure elastic case with an increase in H/E . In figure 4(a), a surface with $Sk = 0.3$ of the three values of Sk considered has

(a)

$$\sigma = 10\text{nm}, \beta^* = 0.8\mu\text{m}, E = 100\text{ GPa}, P_n/E = 10^{-6}, H/E = 0.075$$



(b) Pressure distribution at $z=0$

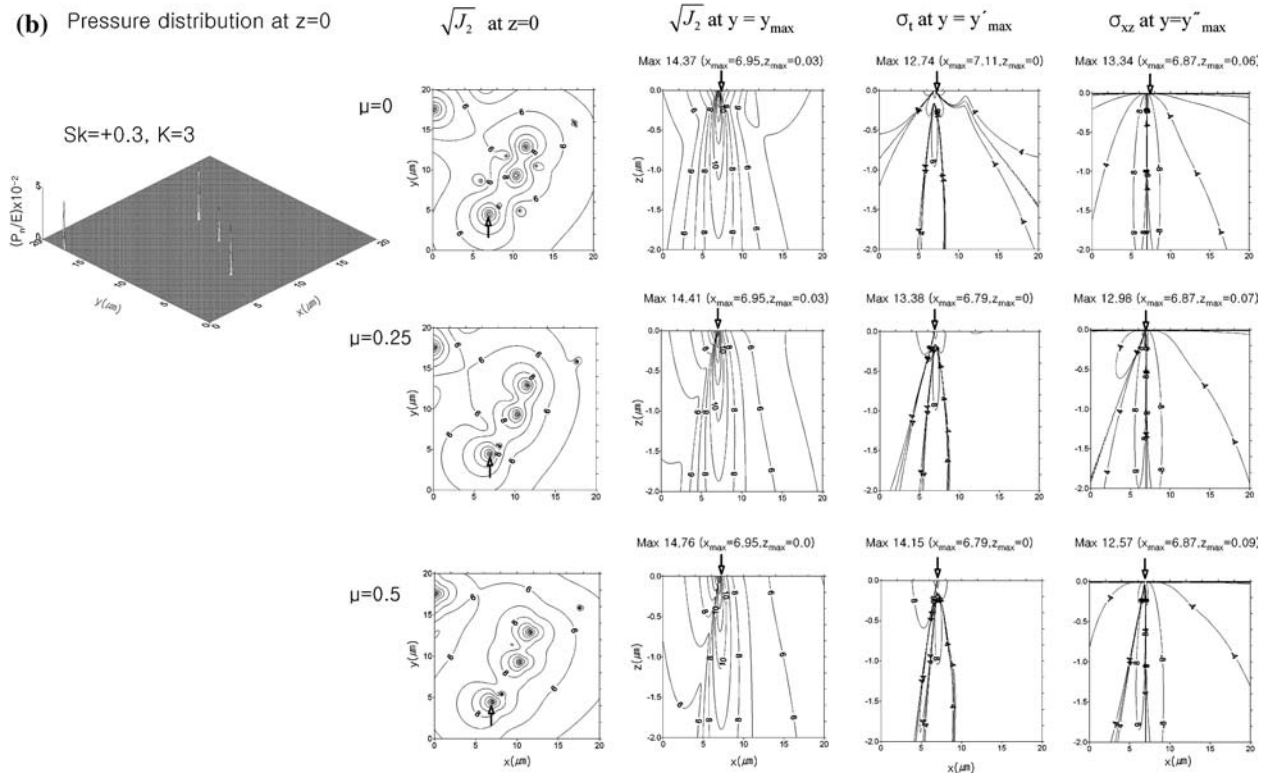


Figure 5. Profiles of contact pressures, contours of von Mises stresses on the surface, von Mises stresses on the max $\sqrt{J_2}$ plane ($y = y_{\max}$), principal tensile stress on the max σ_t plane ($y = y'_{\max}$), and shear stresses on the max σ_{xz} plane ($y = y''_{\max}$) with $\sigma = 10\text{ nm}$, $\beta^* = 0.8\mu\text{m}$, $E = 100\text{ GPa}$, $P_n/E = 10^{-6}$, $H/E = 0.05$ with (a) $Sk = -0.3$, $K = 3$, (b) $Sk = 0.3$, $K = 3$, (c) $Sk = 0$, $K = 3$, (d) $Sk = 0$, $K = 2$, and (e) $Sk = 0$, $K = 3$.

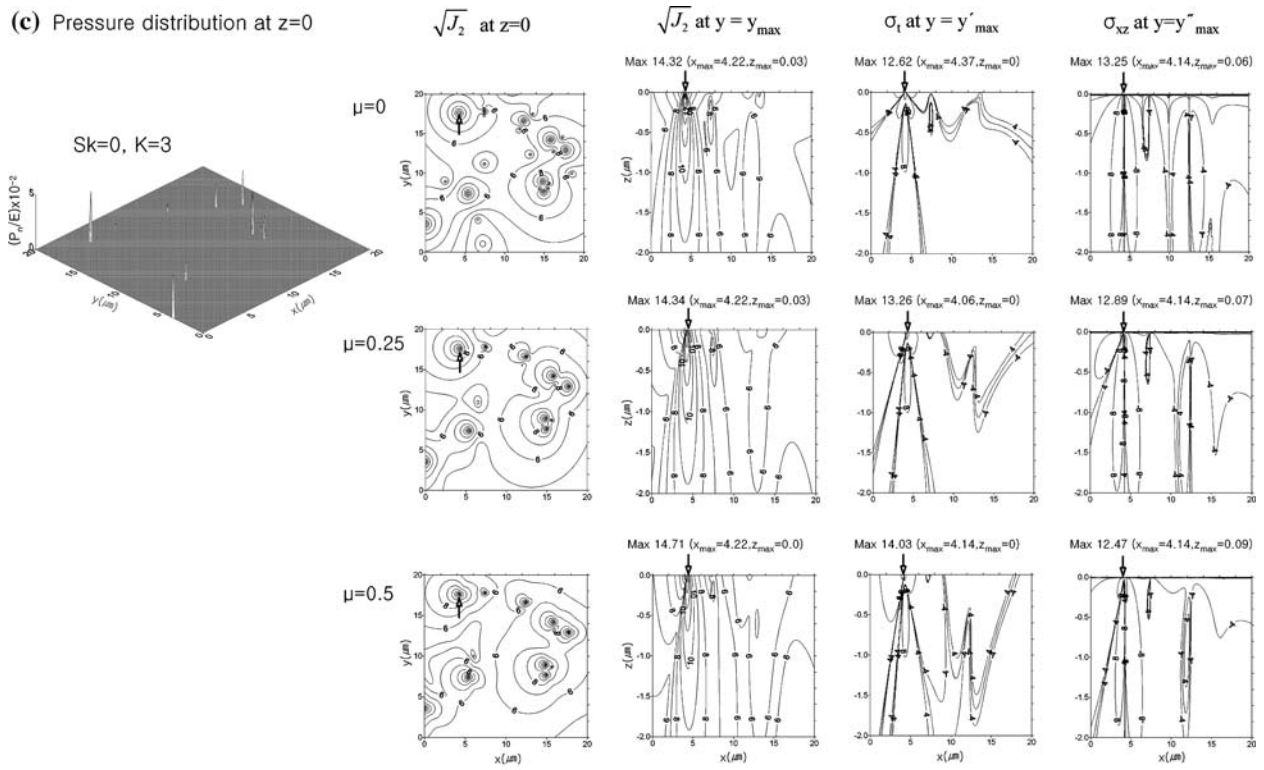
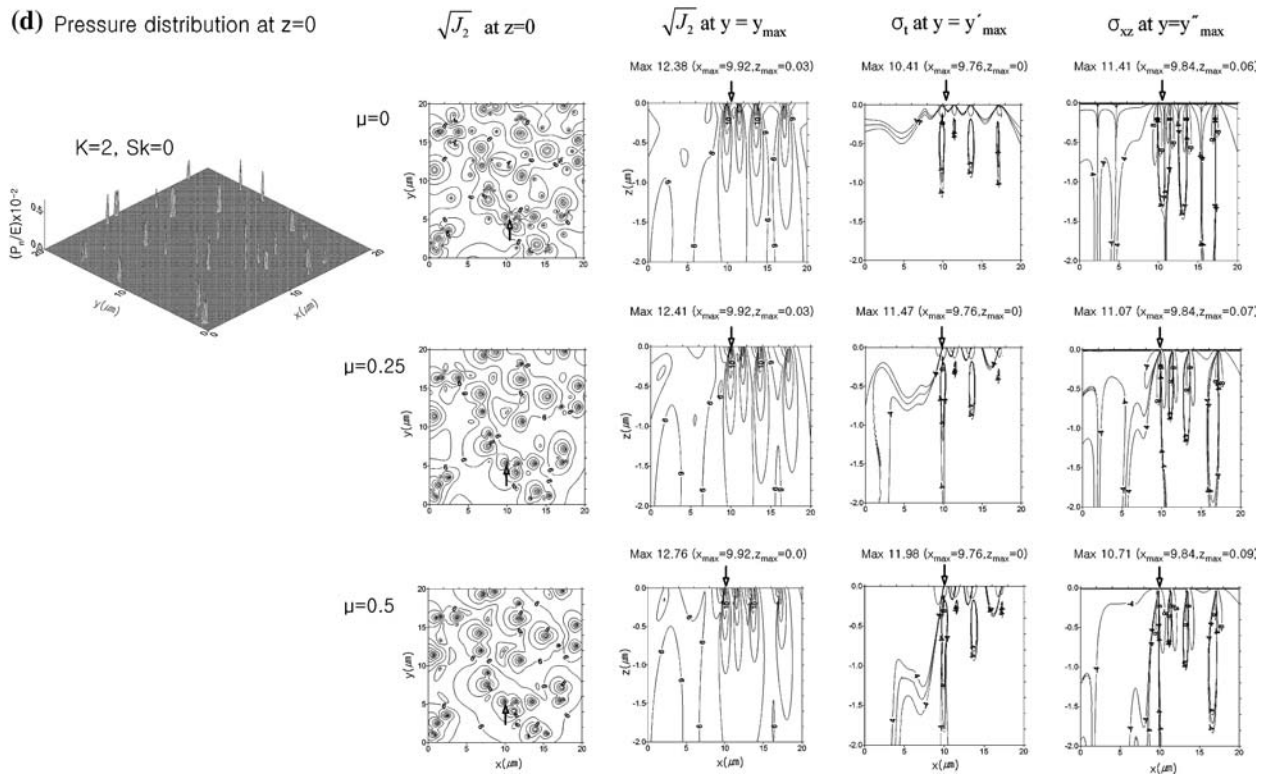
(c) Pressure distribution at $z=0$ (d) Pressure distribution at $z=0$ 

Figure 5. Continued.

the highest critical material hardness. This occurs because a surface with $Sk = 0.3$ gives minimum contact area but high pressures in the contact region, then a ratio of plastic deformed asperities to contacted asperi-

ties is higher than that of other surfaces considered in this study, as shown in the lower figures. Therefore surfaces with $Sk = 0.3$ may fail due to plastic deformation for a surface with low material hardness. For the

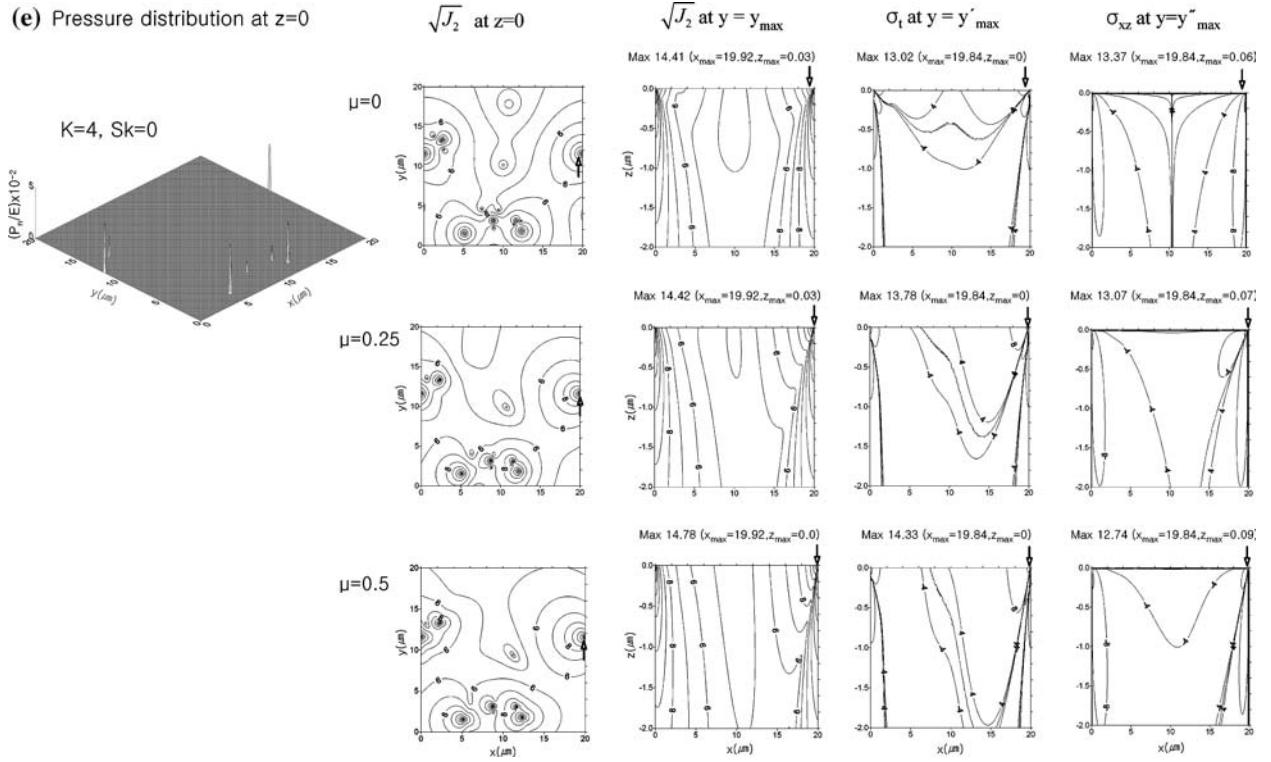


Figure 5. Continued.

effect of kurtosis, the critical material hardness increases with an increase in kurtosis value, due to a decrease in contact area and an increase in plastic area with kurtosis, as shown in figure 4(b). In addition, the critical material hardness of a surface with $K = 2$ is much lower than that of any other surface. This means even though the material of a surface with $K = 2$ has a relatively low hardness, asperities remain an elastic region, which is desirable for low wear. In conclusion, surfaces with $Sk = 0.3$ and $K = 4$ will exhibit a minimum contact area and a maximum interplanar separation which are beneficial to low friction and stiction, but if the hardness of material is not high enough to compensate high contact pressure, high contact pressure may lead to severe plastic deformation.

In a study of the effect of sampling interval on contact pressure [30], the result indicated that an increase of contact stress is coupled with a decrease of sampling interval (an increase of density). In addition, it has been concluded that at very small sampling intervals, contact pressure can exceed the hardness value, leading to plastic deformation. So with a smaller sampling interval than considered here, the results of the critical material hardness may be lower than the present results.

4.3. Subsurface stresses

As stated previously, the patch solution is for elastic analysis. For the surface with a hardness $H/E = 0.075$,

we found that $P_{II}/E = 10^{-6}$ is the highest value that satisfies the condition of no plastic deformation. As shown in the second figures of figure 3, the value of maximum contact pressure/ E in the surface with $Sk = 0$ and $K = 4$ is 0.0745. Therefore, in this section, the applied load is taken as 10^{-6} and the hardness of rough surface is taken as $0.075E$.

Figure 5 shows the pressure distribution on the surface and stress distributions on both the surface and the subsurface with maximum values of stresses and their locations at different skewness and kurtosis values. All contours are plotted after taking natural log values of the calculated stresses expressed in kPa, as the stress has a large range. For the pressure distribution, a surface with $Sk = -0.3$ has a greater number of contact points whereas a surface with positive skewness (0 or 0.3) has higher pressure spikes. The number of contacts decreases and the peaks become higher and sharper with an increase in kurtosis, which results in higher pressure spikes at higher kurtosis values.

The subsurface stress distributions for contact with $\mu = 0, 0.25$ and 0.5 are also shown in figure 5. In a rough surface contact, there is no symmetrical plane to show a typical stress distribution and each individual plane is of unique stress distribution. A $y = \text{constant}$ plane is chosen to make contour plots for subsurface von Mises stress, principal tensile stress and shear stress. Since the locations of maximum $\sqrt{J_2}$, σ_t and σ_{xz} are of interest in the failure prevention, the $y = \text{constant}$ plane is set as the plane including the maximum $\sqrt{J_2}$, σ_t and σ_{xz} , respectively. Concentric contours are observed and

their locations are consistent with those of the contact points as compared with pressure distributions. The contact pressures are seemingly point sources which induce high stresses around the contact points. This implies that under the present applied load, because the contact points are isolated and scattered, the asperity contact behaves like a point load contact. The von Mises stress $\sqrt{J_2}$ is an important factor in predicting an initiation of fatigue crack. As shown in figure, the maximum $\sqrt{J_2}$ always occurs beneath one of the contact asperities with highest contact pressures. As the friction coefficient increases, the subsurface stress contours become asymmetrical and the maximum value grows larger. The location of maximum $\sqrt{J_2}$ has moved to the surface from near surface. Principal tensile stress σ_t makes brittle materials susceptible to ring cracks and is the cause of surface failure. As shown in figure 5, the maximum σ_t always occurs on the surface near a contact point with the maximum pressure. The shear stress σ_{xz} causes the adhesive failure during sliding contact. The

maximum σ_{xz} always occurs beneath the maximum pressure and decays rapidly with the depth [31].

Figure 6 shows the variation of the maximum von Mises and principle tensile stresses for surfaces with different skewness and kurtosis as a function of friction coefficient. As friction coefficient increases, maximum $\sqrt{J_2}$ and maximum σ_t increase for all surfaces. It is interesting to note that the maximum $\sqrt{J_2}$ and maximum σ_t of a surface with $K = 2$ are much lower than any other surface examined. This is due to the distribution of $K = 2$ having relatively few high peaks and low valleys, which results in many contact points and relatively lower contact pressure as shown in figure 5(d). It is beneficial in reducing stress-induced failure but the friction and stiction would be higher.

5. Conclusions

A three-dimensional contact analysis is conducted to investigate the contact behavior of elastic-perfectly

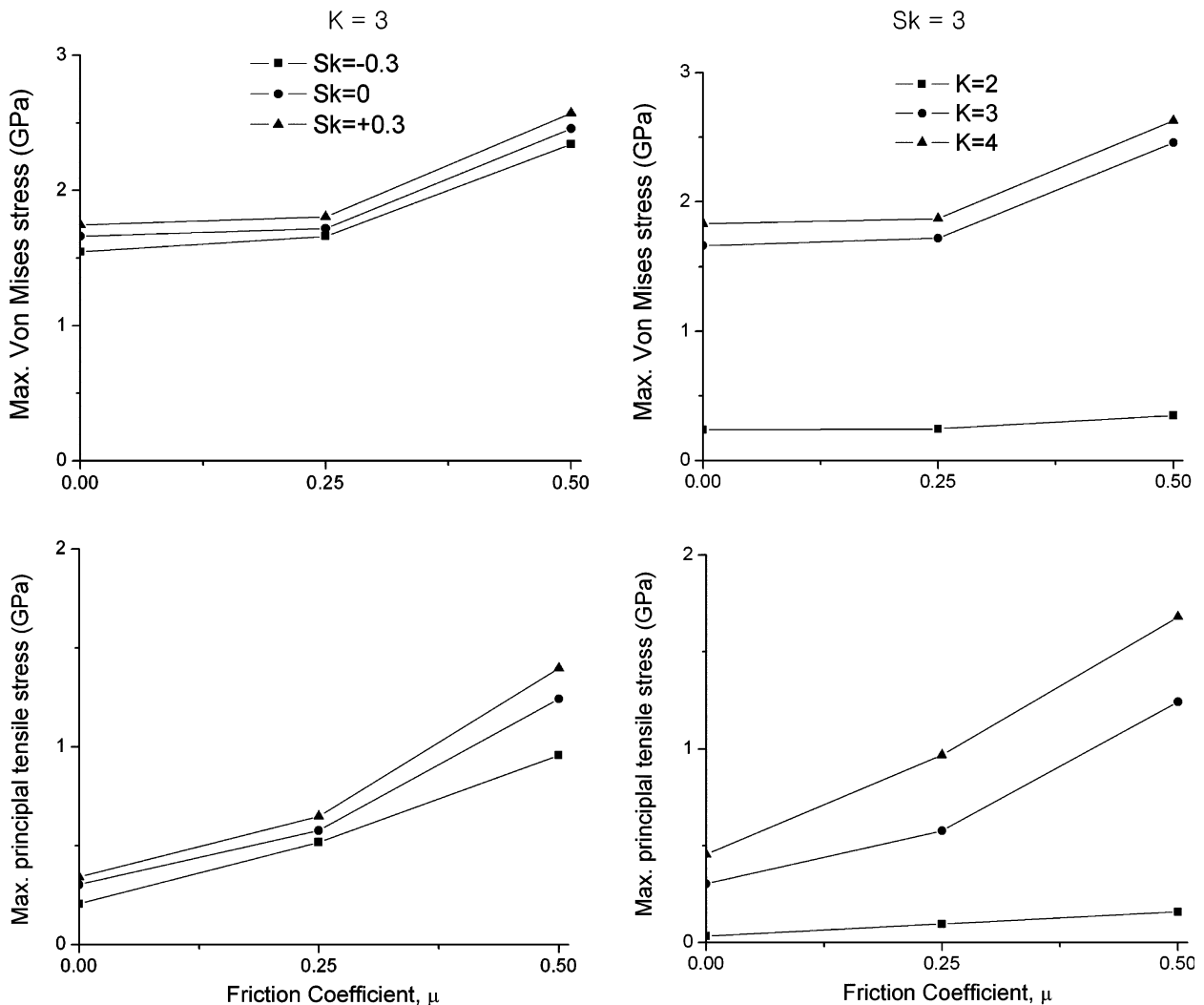


Figure 6. Variation of the maximum von Mises stress and principle tensile for surfaces with different skewness and kurtosis as a function of friction coefficient.

plastic solids with non-Gaussian rough surface. Three-dimensional non-Gaussian rough surfaces are generated on the computer with various skewness and kurtosis values. Contact pressures and subsurface stress fields are obtained by contact analysis of a semi-infinite solid based on the use of influence functions and patch solutions. It is found that for six surfaces with skewness, S_k , of -0.3, 0 and 0.3, and kurtosis, K , of 2, 3 and 4, surfaces with $S_k = 0.3$ and $K = 4$ have a minimum contact area and maximum interplanar separation, which may provide low friction and stiction. The critical material hardness is defined as the hardness at which severe level of plastic asperity deformation corresponding to the Greenwood and Williamson's cut-off $A_{\text{plastic}}/A_{\text{real}} = 0.02$ occurs for a given surface and load condition. The effect of kurtosis and skewness on critical material hardness is investigated. The critical material hardness of surfaces with $S_k = 0.3$ and $K = 4$ is higher than that of other surfaces considered. As the friction coefficient increases, maximum $\sqrt{J_2}$ grows larger and the location of maximum value has moved to the surface from beneath surface; maximum σ_t also increases, but the location is always on the surface. In addition, maximum $\sqrt{J_2}$ and maximum σ_t of a surface with $K = 2$ are very much lower than those of other surfaces considered. This is due to the fact that the distribution of $K = 2$ has relatively few high peaks and low valleys, which results in many contact points and relatively lower contact pressure. It is beneficial in reducing stress-induced failure but the friction and stiction would be higher.

Acknowledgment

This work was supported by the Korea Research Foundation Grant funded by Korea Government (MOEHRD, Basic Research Promotion Fund) (KRF-2004-214-D00213).

References

- [1] J.A. Greenwood and J.B.P. Williamson, Proc. Roy. Soc. A 295 (1966) 300.
- [2] D.J. Whitehouse and J.F. Archard, Proc. Roy. Soc. A 316 (1970) 97.
- [3] R.A. Onions and J.F. Archard, J. of Phys. D. Appl. Phys. 6 (1973) 289.
- [4] M.N. Webster and R.S. Sayles, ASME J. Tribol. 108 (1986) 314.
- [5] X. Tian and B. Bhushan, ASME J. Tribol. 118 (1996) 33.
- [6] B. Bhushan, Tribol. Lett. 4 (1998) 1.
- [7] B. Bhushan and W. Peng, Appl. Mech. Rev. 55 (2002) 435.
- [8] T.R. Thomas, *Rough Surfaces* (Longman, 1982).
- [9] B. Bhushan, *Principles and Applications of Tribology* (John Wiley and Sons, 1999).
- [10] A. Majumdar and B. Bhushan, ASME J. Tribol. 113 (1991) 1.
- [11] B. Bhushan and A. Majumdar, Wear 153 (1992) 53.
- [12] W.T. Lai and H.S. Cheng, ASLE Trans. 28 (1985) 172.
- [13] N. Patir, Wear 47 (1978) 263.
- [14] S.C. Lee and H.S. Cheng, Tribol. Trans. 35 (1992) 523.
- [15] D.J. Whitehouse, *Handbook of Surface Metrology* (Institute of Physics Publishing, 1994).
- [16] T. Merriman and J. Kannel, ASME J. Tribol. 111 (1989) 87.
- [17] R.S. Sayles and T.R. Thomas, Int. Prod. Res. 14 (1976) 641.
- [18] M. Mirazamaani and M.F. Doerner, IEEE Trans. Magn. 32 (1996) 3638.
- [19] C.A. Kotwal and B. Bhushan, Tribol. Trans. 39 (1996) 890.
- [20] S.K. Chilamakuri and B. Bhushan, Proc. Instn. Mech. Engrs. 212 (1998) 19.
- [21] Y.Z. Hu and K. Tonder, Int. J. Mach. Tool Mf. 32 (1992) 82.
- [22] N.L. Johnson, Biometrika 36 (1949) 149.
- [23] W.P. Elderton and N.L. Johnson, *System of Frequency Curves* (Cambridge University Press, 1969).
- [24] M.N. West and R.S. Sayles, ASME J. Tribol. 108 (1986) 314.
- [25] J. Boussinesq, *Application des Potentials a l'Etude de l'Equilibre et du Mouvement des Solides Elastiques* (Gauthier-Villars, 1885).
- [26] D. Tabor, *The Hardness of Metals* (Oxford University Press, 1951).
- [27] A.E.H. Love, Proc. Roy. Soc. Lon. A228 (1929) 377.
- [28] N. Ahmadi, L.M. Keer, T. Mura and V. Vithoontien, ASME J. Tribol. 109 (1987) 627.
- [29] W. Peng and B. Bhushan, ASME J. Tribol. 123 (2001) 330.
- [30] C.Y. Poon and B. Bhushan, J. Appl. Phys. 79 (1996) 5799.
- [31] W. Peng and B. Bhushan, Wear 249 (2001) 741.