Available online at www.sciencedirect.com
ScienceDirect
Wear 261 (2006) 1337–1347
WEAR
www.elsevier.com/locate/wear

A multi-scale model for contact between rough surfaces

Robert L. Jackson [a,*], Jeffrey L. Streator [b]

[a] Department of Mechanical Engineering, Auburn University, Auburn, AL 36849-5341, United States
[b] George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0405, United States

Received 20 September 2005; received in revised form 2 March 2006; accepted 10 March 2006
Available online 2 May 2006

Abstract

This work describes a non-statistical multi-scale model of the normal contact between rough surfaces. The model produces predictions for contact area as a function of contact load, and is compared to the traditional Greenwood and Williamson (GW) and Majumdar and Bhushan (MB) rough surface contact models, which represent single-scale statistical and fractal analyses, respectively. The current model incorporates the effect of asperity deformations at multiple scales into a simple framework for modeling the contact between nominally flat rough surfaces. Similar to the "protuberance upon protuberance" theory proposed by Archard, the model considers the effect of having smaller asperities located on top of larger asperities in repeated fashion with increasing detail down to the limits of current measurement techniques. The parameters describing the surface topography (areal asperity density and asperity radius) are calculated from an FFT performed of the surface profile. Thus, the model considers multi-scale effects, which fractal methods have addressed, while attempting to more accurately incorporate the deformation mechanics into the solution. After the FFT of a real surface is calculated, the computational resources needed for the method are very small. Perhaps surprisingly, the trends produced by this non-statistical multi-scale model are quite similar to those arising from the GW and MB models, but seem largely unaffected by the sampling resolution at the employed surface data.

Keywords: Contact mechanics; Surface roughness; Contact of rough surfaces; Analysis and models

1. Introduction

Since virtually all engineering surfaces are rough to some degree, accounting for the effects of surface topography is critical to modeling surface contact. Effective contact models can lead to an improved understanding of friction and wear, as well as thermal and electrical conductance between surfaces. When two rough surfaces are pressed together, it is primarily the peaks (or asperities) on the surface that will make actual contact. With only a small fraction of the available area supporting the load, the contacting asperities of the surfaces often carry very high compressive stresses. These high stresses will often cause yielding in the material and thus purely elastic contact models of rough surfaces are not always adequate.

One of the earliest models of elastic, rough surface contact is that of Greenwood and Williamson [1]. The Greenwood and Williamson (GW) model applies a contact solution for curved elastic bodies, known as the Hertzian solution [2], to a population of asperities following a given statistical height distribution, while assuming that the deformation of a given asperity is not influenced by the deformation of any other asperity. Various extensions of the GW contact model have been developed to incorporate affects of adhesion and plastic deformation [3–7]. The statistical GW framework has been preserved, but different models are implemented for the asperity deformation. Because the Hertzian contact solution requires the specification of the radius of curvature, this parameter is needed for each of the contacting asperities. In the case of the GW and related models, it is assumed that the asperity radius of curvature is the same for all asperities. In practice, one uses the average radius of curvature of the asperities, as determined from a measurement of the surface profile [8]. One of the drawbacks of this class of models, which rely on the specification of a single radius of curvature, is the ambiguity of scale. That is, the determination of the average radius of curvature of a surface profile is sensitive to the scale of observation, or more specifically, to the lateral resolution used to measure the surface [9–11].

* Corresponding author. Tel.: +1 334 844 3340; fax: +1 334 844 3307.
E-mail addresses: robert.jackson@eng.auburn.edu (R.L. Jackson), jeffrey.streator@me.gatech.edu (J.L. Streator).

0043-1648/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.wear.2006.03.015

Nomenclature

a	radius of the area of contact
A	area of contact
$\bar{A}$	area of contact for single asperity
A_n	nominal contact area of the surface
A_r	real contact area of the surface
B	material dependent exponent
C	critical yield stress coefficient
D	fractal dimension
e_y	yield strength to elastic modulus ratio, S_y/E
E	elastic modulus
E'	$E/(1-\upsilon^2)$
f	spatial frequency (reciprocal of wavelength)
F	contact force
$\bar{F}$	contact force for single asperity
G	fractal roughness parameter
H	hardness
H_G	hardness geometric limit
L	scan length
N	number of asperities
p^*	pressure amplitude
$\bar{p}$	mean pressure
R	radius of hemispherical asperity
S_y	yield strength

Greek symbols

β	asperity amplitude
η	asperity areal density
λ	asperity wavelength
υ	Poisson's ratio
σ	standard deviation of surface heights
ω	interference between hemisphere and surface

Subscripts

c	critical value at onset of plastic deformation
cont	denotes contact
i	frequency level
j	dummy index for frequency level beyond i^*
k	index for frequency level beyond i^*
JG	from model of Jackson and Green [7]
JGH	from model of Johnson et al. [12]

An early appreciation for the influence of multiple scales was demonstrated by Archard [13] whose work precedes the GW model. Archard suggested that the asperities of rough surfaces must be modeled as "protuberance upon protuberance." In other words, any asperity has upon it a collection of smaller asperities, each of which supports a collection of even smaller asperities, and so on. One expects this process to continue down to the atomic scale. With some simplifying assumptions, Archard was able to model this scaling effect for the elastic contact between (1) a rough sphere and a smooth, rigid flat and (2) a nominally flat, rough surface and a smooth, rigid flat. In the latter case, the rough surface was composed of rough hemispherical asperities, having a uniform height distribution. In both cases, Archard showed that as higher and higher levels of refinement are considered (each of which has a characteristic asperity radius of curvature and asperity areal density), the relationship between area of contact and load approached linearity.

This linear relationship between contact load and area provides an explanation as to why Amonton's law of friction appears to hold in many cases. If friction force is linearly proportional to area and area is linearly proportional to load then friction force will be linearly proportional to load (i.e., $F = \mu N$). However, Archard's contact model is based on hypothetical, idealized surfaces and is difficult to apply to a real rough surface. For example, Archard's model does not provide a means of determining the required coefficients from measurements of a surface profile. Archard assumed that each successive level had asperity radii much smaller that in the previous level. Because real surfaces would generally not have successive scales so widely separated, it is problematic to identify such in practice.

A more sophisticated approach to handling the multi-scale nature of surface roughness was offered by Majumdar and Bhushan [14], who developed a fractal based description of surface contact. In that work, the authors assumed a form for the distribution of the contact areas. They equated a given contact diameter to a wavelength and then used the fractal dimension and fractal roughness parameter (as determined from a power spectrum of the surface) to provide an associated asperity amplitude, which they set to be the asperity deformation or interference. Additionally, they used the asperity wavelength and amplitude to provide the radius of curvature of the deformed asperity. With this formulation, Hertzian analysis or the fully plastic truncation model [15] was applied to each contact region and the resulting areas were summed statistically based on the assumed form for the distribution of contact areas. The results provided, among other things, a relationship between the real area of contact and the applied load, which demonstrated power-law behavior.

One criticism of the MB fractal model is that it predicts that the lighter the load, the greater will be the percentage of contacts that are plastically deformed. This counterintuitive result arises from the fact that, for a given surface, the radius of curvature of an asperity is determined solely by the contact diameter, with each contacting asperity always in complete contact. That is, considering an asperity as a half sine wave, the MB model assumes that the asperity is completely flattened. Therefore, there is no provision in the model for gradually increasing the deformation level (i.e., the degree of interference) of an asperity of a given radius of curvature. The authors view this feature as unrealistically restrictive. These faults have also been identified by Borodich [16] in two succinct statements: "the fractal dimension alone cannot characterize the features of contact" and "the problems for elastic rough surfaces cannot be solved using just geometrical arguments and equations of elasticity should be involved."

Other fractal-based models have also been developed by a number of researchers including [17–22]. Persson and coworkers [23,24] also developed a new method for modeling rough surfaces, which is not dependent on fractal roughness, but is not scale dependent like the GW model. However, the model is implemented only for fractal surfaces. Probably most

relevant to the current work is that by Ciavarella et al. [25], which models contact using an idea similar to Archard's and the current work. The major differences is that they model the surface structure between the scales as being fractal and also use a two-dimensional elastic sinusoidal model (Westergaard [26]) instead of the three-dimensional sinusoidal contact and elasto-plastic contact considered in the current work. A subsequent work [27] expands the analysis to provide contact stiffness and resistance. These works also conclude that as higher scales are included in the contact model via fractal mathematics that the contact area will approach zero. This result is not found in the current work, perhaps because the current model is not governed by the fractal description of surfaces.

The present study contributes the understanding of rough surface contact by providing the following: (1) a new, non-fractal-based, multi-scale model that is applicable to elastic as well as elastoplastic conditions, and (2) comparisons to well established statistical and fractal models in the literature.

2. Multi-scale model

2.1. General framework

The current work uses the same direction of thought as Archard [13], but provides a method that can be much more easily applied to real surfaces. The model assumptions, which are somewhat different from those in the GW and MB fractal models, are as follows:

(1) Asperities are arranged so that asperities of smaller cross-sectional surface area are located on top of larger asperities. In the frequency domain this means that asperity distributions of higher frequencies are superimposed upon lower frequency asperities. This is similar to Archard's "protuberance upon protuberance" concept.
(2) Each "level" or frequency of asperities carries the same total load.
(3) The load at each frequency level is shared equally among all the asperities at that level.
(4) At a given frequency level, each asperity deforms according to Hertz theory or to a chosen elasto-plastic asperity contact model, irrespective of the presence of higher frequency asperities upon it.
(5) A given frequency level cannot increase the contact area beyond what is experienced by the frequency level below it.

These assumptions set up the following simple framework of equations for the contact model:

$$A_r = \left(\prod_{i=1}^{i_{max}} \bar{A}_i \eta_i \right) A_n \tag{1}$$

$$F = \bar{F}_i \eta_i A_{i-1} \tag{2}$$

where A_r is the real area of contact, F the contact load, A_n the nominal contact area, and the subscript i denotes a frequency level, with i_{max} denoting the highest frequency level considered. Parameters $\bar{A}_i$ and $\bar{F}_i$ are the single asperity contact area and single asperity contact force at a given frequency level, respectively. The total (nominal) area of contact at a given frequency level is denoted by A_i, while η_i is the corresponding areal asperity density. For example, if a simplified hypothetical case is assumed such that there are only two frequency levels of asperities then Eq. (1) becomes

$$A_r = \bar{A}_2 \eta_2 \bar{A}_1 \eta_1 A_n \tag{3}$$

where $\bar{A}_2$ and $\bar{A}_1$ are the single asperity contact areas and η_1 and η_2 are the areal asperity densities at their respective frequency levels. The product of $\eta_1 A_n$ gives the number of asperities at frequency level 1. Thus the value of $\bar{A}_1 \eta_1 A_n$ gives the contact area of frequency level 1, which is viewed as the nominal contact area from the perspective of level 2. Next, the value of $\eta_2 \bar{A}_1 \eta_1 A_n$ gives the number of asperities at frequency level 2, which, when multiplied by $\bar{A}_2$, the contact area per level 2 asperity, yields the real contact area, A_r. Values for $\bar{A}_1$ and $\bar{A}_2$ are determined from a micro-contact model (e.g., Hertz), assuming that the contact load is equally shared by all asperities of a given level, with the asperity radius of curvature established from the frequency spectrum. In the general case, the repetitive cycle continues until all the asperity frequency levels are considered. Thus, the resulting model uses a recursive approach to predict the real area of contact between two rough surfaces.

To illustrate in more detail how the contact model framework is used to model contact between rough surfaces, a flow chart of the method is given in Fig. 1. After selecting the scan length (L), the input surface data is acquired and an FFT is performed. From the resulting Fourier series, the areal asperity density and radius of curvature are computed for each frequency level according to:

$$\eta_i = 2 f_i^2 \tag{4}$$

$$R_i = \frac{1}{4\pi^2 \beta_i f_i^2} \tag{5}$$

where f_i denotes the frequency (i.e., the reciprocal of wavelength) and β_i is the amplitude corresponding to the given frequency. Although the source of the factor 2 in Eq. (4) may be difficult to see intuitively, the density of asperity peaks is indeed two per reference area, which is defined as one wavelength by one wavelength. This relation can be determined graphically, as done in [12]. The nominal contact area (A_n) is then set to be equal to L^2 and is identified with $i = 0$.

For a given frequency level, the number of asperities is calculated. The total load is then divided evenly among all of the asperities of the given level. Next, the single asperity area of contact is determined from the given asperity load, the given asperity dimensions and the material properties, according to the chosen asperity deformation model (e.g., Hertz). Then multiplying by the number of asperities at that level, a provisional total contact area for that frequency level is computed. In keeping with Assumption #5 above, the result is checked against the contact area predicted by the frequency level below it, and the smaller value is selected as the contact area for the given frequency level. The concept behind this rule is that the frequency level below a

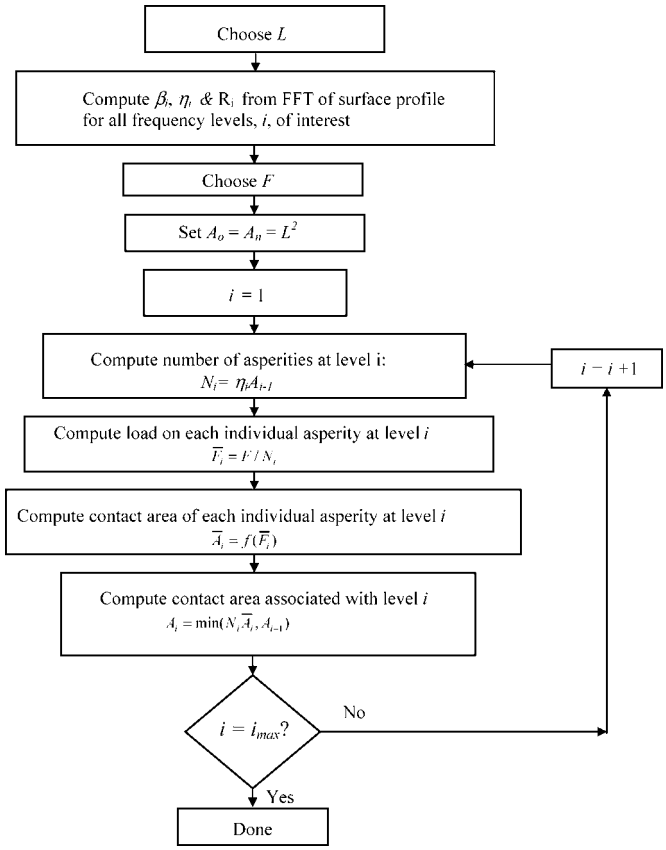

Fig. 1. Flow chart of iterative asperity contact model.

given frequency level serves as the nominal contact area for the given frequency level. The iterative procedure continues until all the desired frequencies are considered, resulting in a prediction of the real area of contact at the prescribed load.

2.2. Asperity deformation

A variety of individual asperity models are available for use within the multi-scale framework described above to relate the contact area to the contact force. For a simulation of purely elastic rough surface contact, the Hertzian model [2] may be applied. Alternatively, several models are available that account for elasto-plastic deformation of spherically shaped asperities [3–7]. A detailed description of several elasto-plastic models is given in Kogut and Etsion [5], and Jackson and Green [7].

Considering again the case of elastic contact, the model framework can also employ the solution to the problem of a sinusoidal surface in contact with a flat, which was solved first by Westergaard [26] for the one-dimensional waviness case and then by Johnson et al. [12] for the two-dimensional waviness case. The analysis for the case of 2D waviness [12] developed by authors Johnson, Greenwood and Higginson (hereafter referred to as "JGH") provides a relation between pressure and contact area. First let $\bar{p}$ be defined as the average pressure in the interface (considering both contacting and non-contacting regions) and let p^* be the amplitude of a sinusoidal pressure superimposed on the mean pressure. Special relations arise when the pressure amplitude p^* is given by:

$$p^* = 2\pi E' \beta f \tag{6}$$

where E' is the reduced elastic modulus, β the waviness amplitude, and f is the reciprocal of wavelength. Namely, when $\bar{p} \geq p^*$, the pressure loads the surfaces so that there is no gap between them. For such a case, the area of contact for the corresponding frequency level is taken to equal that of the frequency level below it, so that the asperities at the given frequency level induce no separation between the surfaces. However, Assumption #5 is upheld no matter which single asperity contact model is used by simply comparing each frequency level to the one just below it. Alternatively, when $\bar{p} < p^*$ the contact is not complete, and a closed form solution for the two-dimensional waviness contact problem is not available. However, Johnson et al. [12] provides two asymptotic solutions to the problem. For $\bar{p} \ll p^*$ the following equation applies:

$$(\bar{A}_{\mathrm{JGH}})_1 = \frac{\pi}{f^2}\left[\frac{3}{8\pi}\frac{\bar{p}}{p^*}\right]^{2/3} \tag{7}$$

and when $\bar{p}$ approaches p^* (i.e., contact is nearly complete) the following equation applies:

$$(\bar{A}_{\mathrm{JGH}})_2 = \frac{1}{2f^2}\left(1 - \frac{3}{2\pi}\left[1 - \frac{\bar{p}}{p^*}\right]\right) \tag{8}$$

Since no general analytical solution is available, we fit a polynomial equation linking Eqs. (7) and (8) based on the experimental data provided by Johnson et al. [12]:

For $\frac{\bar{p}}{p^*} < 0.8$:

$$\bar{A} = (\bar{A}_{\mathrm{JGH}})_1\left(1 - \left[\frac{\bar{p}}{p^*}\right]^{1.51}\right) + (\bar{A}_{\mathrm{JGH}})_2\left(\frac{\bar{p}}{p^*}\right)^{1.04} \tag{9}$$

For $\frac{\bar{p}}{p^*} \geq 0.8$:

$$\bar{A} = (\bar{A}_{\mathrm{JGH}})_2 \tag{10}$$

The curve fit is shown in Fig. 2 with the asymptotic solutions from Eqs. (7) and (8) as well as the data from a numerical sim-

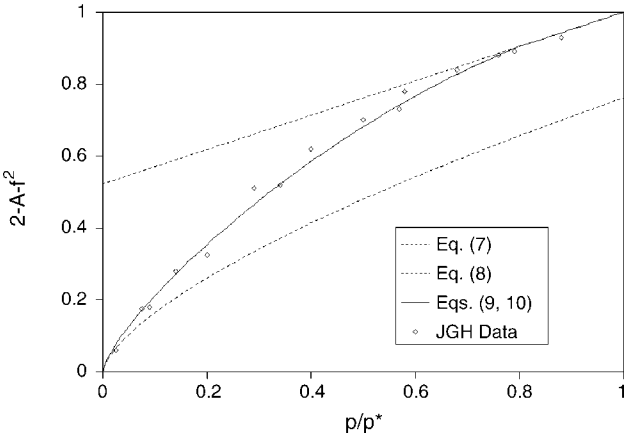

Fig. 2. Real area of contact of two-dimensional wavy surface.

ulation and experimental measurements obtained from [12]. As observed, the curve fit faithfully captures the trend of the data.

As the load pressing two rough surfaces together increases, the stresses within the individual asperities also increase. These stresses eventually cause the material within the asperity (modeled as a hemisphere) to yield. The interference at this initial point of yielding is known as the critical interference, ω_c. Jackson and Green [7] and Chang et al. [3] derive this critical interference analytically using the von Mises yield criterion. The resulting equation [7] is:

$$\omega_c = \left(\frac{\pi C S_y}{2E'}\right)^2 R \tag{11}$$

where C is given by

$$C = 1.295\exp(0.736\nu) \tag{12}$$

and ν is the Poisson ratio of the material that yields first. An alternative to Eq. (11), which is explicitly dependent on material hardness, is given by [3].

The critical force, $\bar{F}_c$, is then calculated at the critical interference, ω_c, to be

$$\bar{F}_c = \frac{4}{3}\left(\frac{R}{E'}\right)^2\left(\frac{C}{2}\pi S_y\right)^3 \tag{13}$$

Similarly, the critical contact area is

$$\bar{A}_c = \pi^3\left(\frac{C S_y R}{2E'}\right)^2 \tag{14}$$

Notice that Eqs. (11)–(14) do not require specification of the hardness (i.e., the presumed maximum achievable surface pressure), which has recently been shown to depend on the degree of indentation [7,28,29], and be influenced by strain gradient effects [30–36]. The current multi-scale contact model is capable of incorporating such effects via selection of an appropriate single-asperity deformation model.

In the present study, the multi-scale framework (see Section 2.1) is used in combination with either Hertzian or the JGH analysis at the asperity level to model elastic contact; whereas to model elasto-plastic contact, the multi-scale framework is used in conjunction with either the asperity deformation model of Jackson and Green ("JG") [7] or the truncation model [14]. While these are not the only modeling options [3,5,6], their use will suffice for investigating the multi-scale formulation.

The JG model predicts the contact force and area between an elastic, perfectly plastic hemisphere and a rigid flat as functions of the ratio of interference (ω) to critical interference (ω_c). (The converse case of a rigid sphere indenting an elastic-perfectly plastic half-space is not considered here.) At $0 \le \omega/\omega_c \le 1.9$ the JG single asperity model virtually coincides with the Hertzian solution. At interference ratios (ω/ω_c) larger than 1.9 the formulation below is used in the JG model for single-asperity elasto-plastic contact:

For $\omega \ge 1.9\omega_c$

$$\bar{A}_{JG} = \pi R\omega\left(\frac{\omega}{1.9\omega_c}\right)^B \tag{15}$$

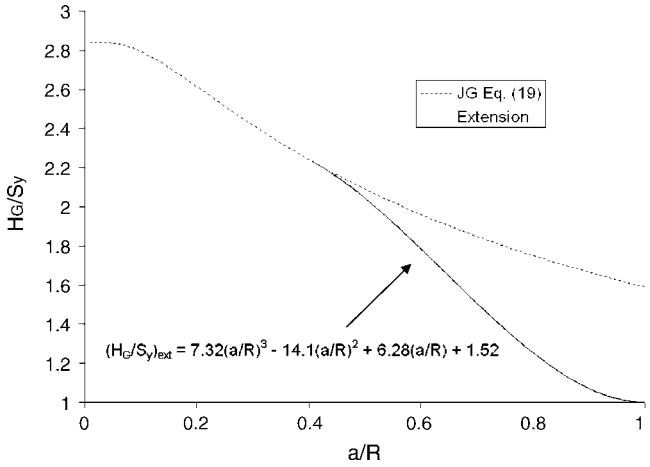

Fig. 3. Extension of JG geometrically varying hardness empirical model for heavy deformations ($a/R > 0.41$).

$$\bar{F} = \bar{F}_c\left\{\left[\exp\left(-\frac{1}{4}\left(\frac{\omega}{\omega_c}\right)^{5/12}\right)\right]\left(\frac{\omega}{\omega_c}\right)^{3/2} + \frac{4H_G}{CS_y}\left[1-\exp\left(-\frac{1}{25}\left(\frac{\omega}{\omega_c}\right)^{5/9}\right)\right]\frac{\omega}{\omega_c}\right\} \tag{16}$$

where

$$B = 0.14\exp(23e_y) \tag{17}$$

$$e_y = \frac{S_y}{E} \tag{18}$$

$$\frac{H_G}{S_y} = 2.84\left[1-\exp\left(-0.82\left(\frac{a}{R}\right)^{-0.7}\right)\right] \tag{19}$$

The parameter H_G, referred to as the "hardness geometric limit" [7], represents an asymptote for mean compressive stress, for increasing contact radius. Quicksall et al. [37] confirmed these results for a range of materials by varying E, S_y and ν. It should be noted that the validity of Eq. (19) is not established for a/R values greater than 0.41 [7] and, further, the expression on the right-hand side returns unrealistic values as a/R approaches unity. Yet, in using the current contact model, the value of a/R can be much larger than 0.41 for a given frequency level, depending on the degree of loading and the details of the frequency spectrum, Similarly, in the context of a GW-type statistical formulation, the higher asperities may experience large enough values of interference for a/R to exceed 0.41. To handle such cases, a polynomial fit is used to extend beyond the valid range of Eq. (19). The coefficients of the polynomial are found by satisfying the following four boundary conditions: (1) the polynomial matches Eq. (19) at $a/R = 0.41$, (2) the polynomial slope matches the slope of Eq. (19) at $a/R = 0.41$, (3) the polynomial equals one, or the limiting value of H_G/S_y at $a/R = 1$ [7] and (4) the slope of the polynomial equals zero at $a/R = 1$. This matching process results in the polynomial extension of Eq. (19) that is shown in Fig. 3. Once the asperity has deformed to the extent that $a/R = 1$, the surfaces are considered to be in complete contact. This extension depicted

in of Fig. 3 is used in conjunction with the JG model for the remainder of this work.

One limitation of the JG model is that for very large deformations (approximately $\omega > 5000\omega_c$) the predicted average normalized pressure, $\bar{F}/(\bar{A}S_y)$, can reduce to a value less than one, which is not physically possible while the asperity is deforming elasto-plastically. This unphysical result occurs because the JG model was formulated based on FEM data developed for much smaller deformations. For this reason and for lack of a better model, the minimum value of $\bar{F}$ is simply set to $\bar{A}S_y$ whenever Eq. (16) returns a value less than this quantity.

2.3. Power-law roughness with Hertzian asperity contact

In the case of Hertzian asperity contact, the relations (Eqs. (1) and (2)) can be used to derive analytical relationships for the contact area. Letting $\bar{A}_i$ denote the single-asperity contact area at frequency level i, we have that

$$\bar{A}_i = \pi a_i^2 = \pi\left(\frac{3\bar{F}_i R}{4E'}\right)^{2/3} \tag{20}$$

where $\bar{F}_i$ is the load supported by the asperity, which has a contact radius a_i. Using Eq. (5) in (20) yields

$$\bar{A}_i = \pi\left(\frac{3\bar{F}_i}{8\pi^2 E'\beta_i\eta_i}\right)^{2/3} \tag{21}$$

Letting N_i denote the number of asperities at frequency level i, then we have $\bar{F}_i = F/N_i$ and $\bar{A}_i = A_i/N_i$. But, by definition, $N_i = \eta_i A_{i-1}$, so that

$$\bar{A}_i = A_i/\eta_i A_{i-1} \tag{23}$$

Similarly, we have

$$\bar{F}_i = F/\eta_i A_{i-1} \tag{24}$$

which could have also been obtain directly from Eq. (2). For convenience, Eq. (24) can be manipulated to give

$$\bar{F}_i = \left(\frac{F}{A_n}\right)\left(\frac{A_n}{A_{i-1}}\right)\frac{1}{\eta_i} \tag{25}$$

Substituting the relations (23), (25) and (4) in (21) gives rise to

$$\frac{A_i}{A_n} = \frac{1}{4}\left(\frac{9}{2\pi}\right)^{1/3}\left(\frac{F}{E'A_n}\right)^{2/3}\left(\frac{1}{f_i\beta_i}\right)^{2/3}\left(\frac{A_{i-1}}{A_n}\right)^{1/3} \tag{26}$$

The above equation represents a recursion relation that expresses the contact area corresponding to a given frequency level in terms of contact area associated with the frequency level below the given one. In the context of the current model, Eq. (26) is valid provided that the predicted contact area at the given frequency level does not exceed that of the previous level (Assumption #5).

It is instructive to consider two special cases of a surface topography. First consider a frequency spectrum that satisfies $\beta_i = b/f_i$, for all i greater than some frequency level i^*, where b is a dimensionless constant (i.e., independent of i). Physically, this case corresponds to the amplitude decaying according to the reciprocal of frequency. For this type of spectrum, (26) can be written as

$$\frac{A_i}{A_n} = C\left(\frac{A_{i-1}}{A_n}\right)^{1/3} \quad (i > i^*) \tag{27}$$

where

$$C = \frac{1}{4}\left(\frac{9}{2\pi}\right)^{1/3}\left(\frac{F}{E'A_n b}\right)^{2/3} \tag{28}$$

The solution of (27) is given by

$$\frac{A_{i^*+k}}{A_n} = \left(\frac{A_{i^*}}{A_n}\right)^{1/3^k}\prod_{j=1}^{k} C^{1/3^{j-1}} = \left(\frac{A_{i^*}}{A_n}\right)^{1/3^k} C^{3/2(1-(1/3^k))} \tag{29}$$

Now Eq. (29) expresses the contact area corresponding to any frequency level beyond i^*, up to the highest frequency level under consideration. It can be readily shown that A_{i^*+k} as given by (29) will be a monotonically decreasing function of k provided that $C^{3/2} < A_{i^*}/A_n$. In this case (29) will be in accordance with Assumption #5 all the way to the highest frequency level of the surface profile. Assuming $k \gg 1$ it is seen from (29) that the first term on the far right-hand side represents a quantity being raised to a very small power. Even when the ratio A_{i^*}/A_n is small the resulting expression tends toward unity. For example, with $A_{i^*}/A_n = 10^{-4}$ and $k = 10$, $(A_{i^*}/A_n)^{1/3^k} = (10^{-4})^{1/3^{10}} = 0.99984$. So for large k, Eq. (29) simplifies to $A_r/A_n = C^{3/2}$, where A_r is the real area of contact. Now using the definition of C from Eq. (28), we obtain

$$A_r = \frac{3}{8\sqrt{2\pi}}\frac{F}{E'b} \tag{30}$$

Therefore, in this very special case, we have a precise linear dependence of real area of contact on load.

Another, more general special case, occurs when, beyond a certain frequency level we have that $\beta_i = b_m f_i^{-m}$, where, b_m is a constant and m is a positive number. Now, by definition of spatial frequency, $f_i = i/L$, where L is the longest wavelength of the profile, which is taken to be equal to $\sqrt{A_n}$. Now defining a dimensionless constant b as b_m/L^{1-m}, one can write $\beta_i = Lbi^{-m}$, so that $\beta_i f_i = bi^{1-m}$. Using this relation in (26) yields

$$\frac{A_i}{A_n} = Ci^p\left(\frac{A_{i-1}}{A_n}\right)^{1/3} \quad (i > i^*) \tag{31}$$

where $p = 2(m-1)/3$. The solution of (31) is given by

$$\frac{A_{i^*+k}}{A_n} = \left(\frac{A_{i^*}}{A_n}\right)^{1/3^k} C^{3/2(1-(1/3^k))}\prod_{j=1}^{k}(i^*+j)^{p/3^{k-j}} \tag{32}$$

It is seen immediately that Eq. (32) reduces to the previous special case considered when $p = 0$ (i.e., $m = 1$). Therefore, for k large, the product of the first two terms on the right-hand side tends rapidly toward $C^{3/2}$. On the other hand, it is readily shown that for $p > 0$ the product term is a monotonically increasing function of k. Hence at some value of k, A_{i^*+k} must itself become a monotonically increasing function of k. Therefore, depending

primarily on the value of i^*, we expect to reach a frequency level $i = i^* + k$ where A_i represents the real area of contact. That is, given the restriction of Assumption #5, beyond this frequency level Eq. (32) would no longer apply and the most recent calculated area (i.e., at the given frequency level) would serve as the real area of contact. While it is not apparent from inspection of (32) what is the frequency level at which the right-hand side becomes an increasing function of k, it is plausible that the area corresponding to this frequency level is somewhat insensitive to the load: for large k, the ratio of consecutive areas would be virtually independent of load because they would both contain the factor $C^{3/2}$, which would cancel out (see Eq. (28)). Thus, only for small k, would there be any possible dependence on load. But, since the right-hand side of (32) has only weak dependence on k, small changes in the critical k value (i.e., where the function becomes increasing) due to changes in load (and therefore C) would have only a small impact on the contact area given by (32). Hence, over a wide range of loads, it is anticipated that the real area of contact will be nearly proportional to $C^{3/2}$, which implies linearity with load.

3. Results and discussion

3.1. Basis of comparison

In this study, the new multi-scale model is compared with other models for both elastic and elasto-plastic contact. For the case of purely elastic contact, the multi-scale framework incorporates the Hertzian or JGH analysis at the asperity level and is compared to the GW model (which uses Hertzian analysis). For the case of elasto-plastic contact, two types of comparisons are made: First the multi-scale framework incorporates the JG analysis at the asperity level and is compared to a modified GW statistical model that also incorporates the JG asperity deformation model. Secondly, the multi-scale framework incorporates a Hertzian analysis in conjunction with a (fully plastic) truncation model [14] at the asperity level, and is compared to the MB fractal model, which similarly employs Hertzian analysis for elastically deformed asperities, while using the truncation model for those asperities that have exceeded the elastic limit. Note, in both cases involving the truncation model, there is a discontinuity in the single-asperity contact area at the transition from elastic to fully plastic contact. For the present study, the fractal analysis follows the methodology of references [9,14] wherein equations are provided expressing load and contact area as functions of the fractal parameters D and G, as well as the material properties of the surface.

If the plastic deformation covers the entire area of contact, it is said that a fully plastic condition is reached. The fully plastic truncation model states that under fully plastic conditions the area of contact of an asperity pressed against a rigid flat can be approximately calculated by mathematically truncating the asperity tip as the rigid flat translates an interference, ω, into the tip. Then the average pressure between the asperity and the flat is assumed to be equal to hardness, or approximately three times the yield strength ($H = 2.84S_y$ is used in this work). However, FEM results [7] show that this model oversimplifies the picture a bit, in that the "hardness" if taken to be the average contact pressure, can actually decrease with increasing interference. Although the truncation model is often attributed to Abbott and Firestone [15], they actually intended their model to be used to describe a wear process rather than an indentation process.

3.2. Computational results

An experimentally obtained surface profile is used to compare the new multi-scale model to statistical (i.e., GW-type) and fractal formulations. A stylus profilometer was used to measure this profile from an arbitrary machined metal sample and is shown in Fig. 4a. The displayed profile is comprised of 2454 data points, so that the horizontal resolution is 0.204 μm. Although more detailed profiles might be obtained using higher resolution equipment (e.g., an AFM), since all the models will use the same surface data, the given measurements are adequate for making a comparison among the models considered in this paper. Moreover, the given resolution is likely representative of what might

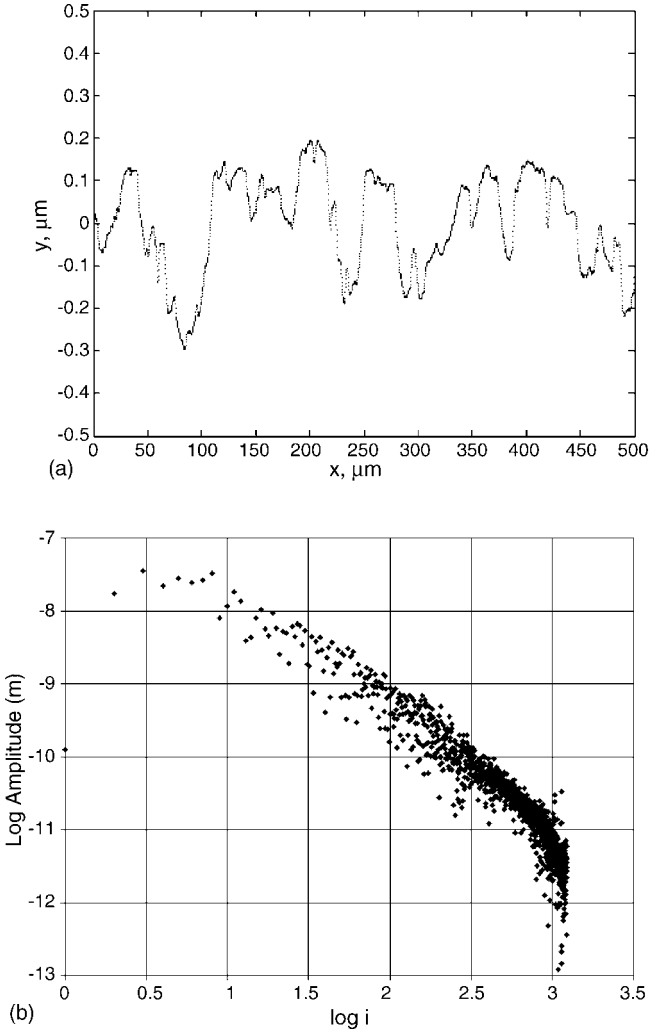

Fig. 4. Surface profile from a stylus surface profilometer (a) and the resulting frequency spectrum of the surface (b).

Table 1
Statistical, fractal and material parameters used in current work

Category	Parameter	Name	Value
Statistical	η	Asperity areal density	$3.880 \times 10^{11}/m^2$
	σ	rms roughness	0.115 μm
	R	Asperity radius of curvature	9.299 μm
Fractal	D	Fractal dimension	1.597
	G	Fractal roughness parameter	2.86×10^{-10} m
Material properties	E'	Elastic modulus	200 GPa
	S_y	Yield strength	0.2 GPa
	ν	Poisson's ratio	0.32

be obtained in an industrial setting. The same material properties are used for all simulations and are given in Table 1.

An FFT of the surface profile is shown in Fig. 4b. At the higher frequencies, the computed amplitudes, being much less than an angstrom are sufficiently small so as to lose some physical meaning. Most of the real surface data are reflected at frequency levels corresponding to i less than 100. Now the profile's autocorrelation function produces the surface's power density function. A power density function was fit to the surface data to find the fractal parameters D and G using the methods outlined in Majumdar and Tien [10]. However, based on the range of frequencies considered when fitting the PDF, the predicted values of D and G can vary significantly for the same surface. This is especially true of a real surface, which may not be a true fractal surface. For this reason, the results of the various models are also compared using a generated fractal surface in a later section. The methods derived by McCool [8] are used to extract the statistical parameters (i.e., σ, η, R) from the profile by finding the spectral moments of the surface. The resulting fractal and statistical parameters used in this work are given in Table 1. It is noted here that the statistical parameters are dependent upon the sampling resolution. For example, if we re-compute the statistical parameters based on a resolution of 0.4 μm, then the resulting values for σ, η, R are 0.0815 μm, $1.20 \times 10^{11}/m^2$ and 25.0 μm, respectively. Comparing with the values in Table 1, we see that both R and η are quite sensitive to the choice of resolution. Such a dependency is expected, as it has been shown previously that the choice of sampling resolution can bias the statistical parameters [9–11].

To further emphasize how the current model works, the contact area predicted by the model as a function of the frequency level iteration for the case of $F/(A_n E') = 7.94 \times 10^{-5}$ is shown in Fig. 5. Segments where the values decrease indicate frequency levels that cause reduction in the real area of contact relative to the previous level. It is clear that a few lower frequency ranges dominate the calculations and the higher frequencies seem to have very little effect on the predicted contact area. For a wide range of loads it also appears that these same lower frequency ranges dictate the results. This is in contrast to in the predictions of Ciavarella et al. [25], which suggest that the contact area tends toward zero as higher and higher frequencies are included.

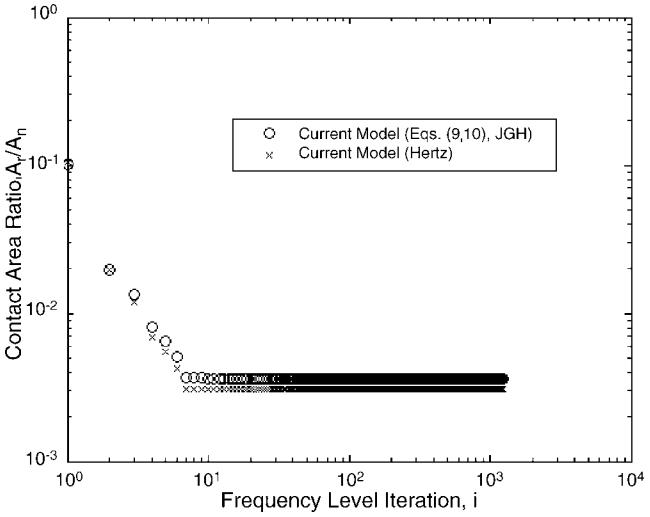

Fig. 5. Decrease in predicted contact area as a function of considered frequency levels.

The results produced by the current multi-scale model using Hertzian or JGH analysis (Eqs. (9) and (10)) at the asperities (both of which are purely elastic models) are shown in Fig. 6 alongside the classic GW model, which assumes a normal distribution of asperity heights. The current model predicts a larger non-dimensional contact area as a function of the non-dimensional surface load than the GW model. In addition, the current model using Eqs. (9) and (10) predicts a slightly larger contact area than the current model using Hertzian contact. The GW curve reflects a near linear relationship between area and load, with a slope of 0.969 in the log–log domain. Both applications of the current multi-scale model show almost perfect linearity with slopes of 1.00 in the log–log domain. Note that a slope of unity reflects a direct proportionality between load and area.

Fig. 7 shows the results for the various elasto-plastic contact models described above, including the current framework using the Jackson and Green formulation (identified as "JG") [7] and

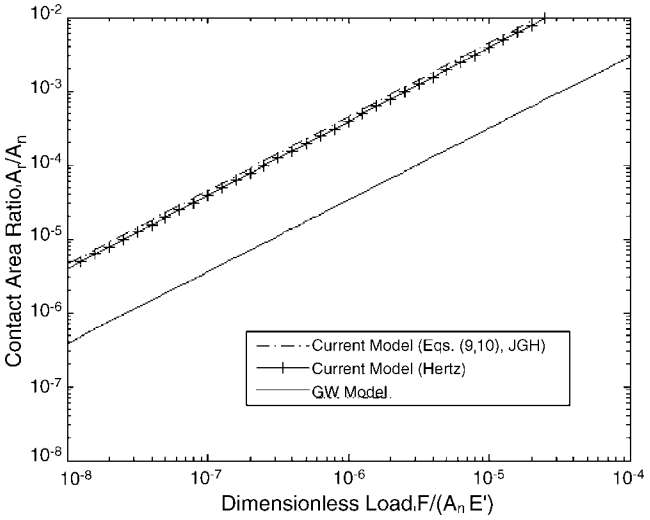

Fig. 6. Comparison of elastic rough surface contact models.

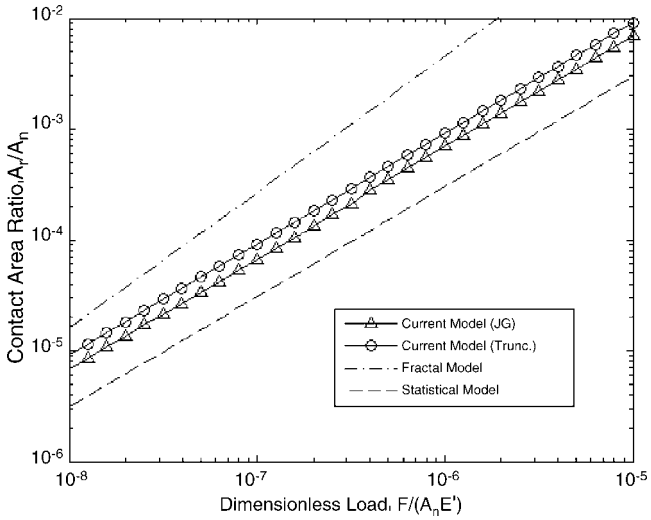

Fig. 7. Comparison of elasto-plastic surface contact models.

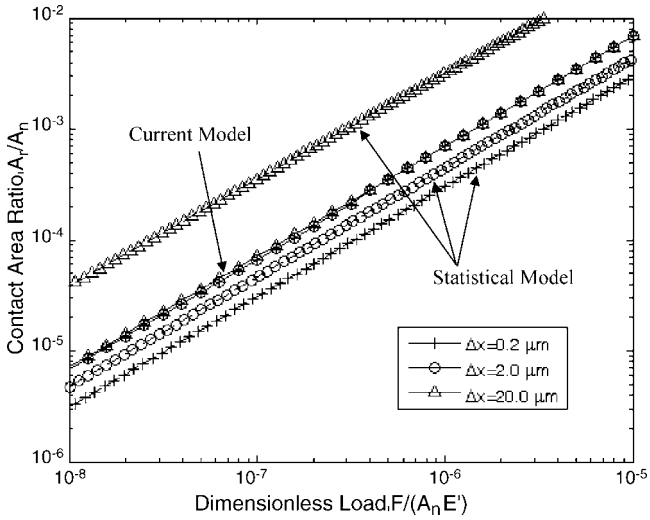

Fig. 8. Comparison of current and statistical models for various horizontal resolutions (Δx).

also, the current framework using a combination of Hertz contact [38] for when the asperity is deforming elastically and the truncation model [15] when the asperity is deforming elasto-plastically (identified by "Trunc."). Also shown are results for the Majumdar and Bhushan (MB) fractal model [14], and the GW statistical contact model using the JG formulation. Interestingly, both applications of the current model predict surface contact behavior that locates between the MB fractal model and the statistical model.

In Fig. 7, the slope of the curve corresponding to the current framework, applying either the JG or truncation asperity deformation models is nearly unity as was the case in Fig. 6, where only elastic deformation was considered. So, again we find proportionality between area and load. This proportionality has an interesting physical interpretation: The average contact pressure based on the *real* contact area is given by

$$\bar{p}_{\text{cont}} = F/A_{\text{r}} \tag{33}$$

which can be written as:

$$\bar{p}_{\text{cont}} = [(F/A_{\text{n}}E')/(A_{\text{r}}/A_{\text{n}})]E' \tag{34}$$

The term in brackets gives the ratio of the normalized load to the normalized real area of contact. When the real area of contact is proportional to the load, this ratio is a constant, suggesting that there is a characteristic asperity contact pressure that is independent of the load. For example, in Fig. 6, the Hertz case corresponds to a characteristic contact pressure of $0.00255 \times E'$. Similar considerations apply when the multi-scale model incorporates elastic–plastic contact as in Fig. 7. Therefore, the multi-scale model suggests that the mean contact pressure is dictated by the details of the surface topography and the elastic constants and yield strength, as opposed to the external load. Increasing the load serves to bring more asperities into contact, but does not change the load supported per contact area. It is left to a future study to explore how the specifics of the surface topography influence the characteristic pressure.

Predictions using several horizontal resolutions (Δx) are displayed in Fig. 8 to illustrate the insensitivity of the current model to the choice of sampling resolution (Δx)—in contrast to the GW statistical type models. As observed, the predictions of the current model show only a minute change when the sampling resolution is increased over two orders of magnitude from the reference value of 0.2 μm (which was used in Fig. 7) By using a larger Δx, higher frequencies are effectively neglected, but as shown in Fig. 5 and discussed in Section 2.3, the current model does not appear to be influenced by the higher frequencies. The results with the current model are in drastic contrast to the GW statistical type models, which are significantly affected by Δx.

The fractal parameters, G and D are obtained through a fit of the power density spectrum. Because the fit is not perfect, the fractal characterizations are an approximation of the surface profile. As a result, the values of G and D are largely dependent on the frequency range used to calculate them for surfaces that are not true fractals. In order to apply the current model to a true fractal (and thereby make a fairer comparison with a fractal model) a true fractal surface is generated using the equations given by Majumdar and Bhushan [14] at a resolution of 1.0 μm, using the fractal parameter values given in Table 1. The current multi-scale model is then applied to the generated fractal surface and compared to the predictions of the MB fractal model. To compare both of these predictions to those of the GW model, McCool's methods were used to extract the GW parameters from the fractal surface. The highest resolution available from the surface data (1.0 μm) was used to derive the surface parameters from the spectral moments The results are shown in Fig. 9 for the elasto-plastic fractal model (MB), statistical model and current model. The fractal results in Figs. 7 and 9 are identical because the same fractal parameters are used. The results for the current model are relatively unchanged from the previous results (Fig. 7), while the statistical model curve has translated upward significantly and now intersects with the fractal model curve (although the slopes are quite different).

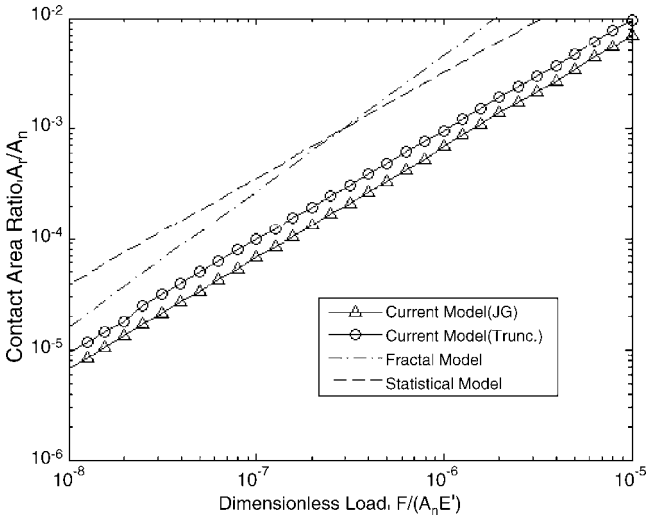

Fig. 9. Comparison of elastoplastic contact models derived from generated fractal surface using fractal parameters in Table 1.

4. Conclusions

This work examined an iterative multi-scale framework for modeling contact between rough surfaces, which uses the Fourier series coefficients resulting from a FFT of the surfaces to characterize the surface geometry at multiple scales. This model attempts to capture multi-scale characteristics not considered in the Greenwood and Williamson model [1] while providing a more accurate account of asperity deformation than found in the Majumdar and Bhushan [14] fractal model. Additionally, the current model is able to consider multiple scales without being restricted to fractal surfaces. The multi-scale framework allows for the introduction of most any elastic, elasto-plastic or fully plastic asperity deformation model.

The multi-scale model was applied to a sample surface profile. Several conclusions can be made as a result of the calculations:

(1) Only the lower frequencies impact the real area of contact.
(2) The real area of contact is proportional to the load for both elastic and elasto-plastic contact.
(3) The mean asperity contact pressure is independent of the load over a wide range loads.
(4) The results of the multi-scale model are insensitive to the sampling resolution.

The limitations of the presented model need to also be mentioned. Probably the most significant weakness is that the model assumes that all asperities at a given frequency behave identically in terms of deformation, load support, etc. Thus there is no provision for the fact that asperities of a given wavelength may reside at different heights. Also, if a single asperity contact model is used, interaction between the laterally spaced asperities is not considered. This shortcoming, however, is overcome by using a sinusoidal contact model as demonstrated in the current work.

References

[1] J.A. Greenwood, J.B.P. Williamson, Contact of nominally flat surfaces, Proc. R. Soc. Lond. A 295 (1966) 300–319.
[2] S. Timoshenko, J.N. Goodier, Theory of Elasticity, McGraw-Hill, New York, 1951.
[3] W.R. Chang, I. Etsion, D.B. Bogy, An elastic–plastic model for the contact of rough surfaces, ASME J. Tribol. 109 (2) (1987) 257–263.
[4] A. Polycarpou, I. Etsion, Analytical approximations in modeling contacting rough surfaces, ASME J. Tribol. 121 (2) (1999) 234–239.
[5] L. Kogut, I. Etsion, Elastic–plastic contact analysis of a sphere and a rigid flat, J. Appl. Mech., Trans. ASME 69 (5) (2002) 657–662.
[6] Y. Zhao, D.M. Maletta, L. Chang, An asperity microcontact model incorporating the transition from elastic deformation to fully plastic flow, ASME J. Tribol. 122 (1) (2000) 86–93.
[7] R.L. Jackson, I. Green, A finite element study of elasto-plastic hemispherical contact, ASME J. Tribol. 127 (2) (2005) 343–354.
[8] J.I. McCool, Comparison of models for the contact of rough surfaces, Wear 107 (1) (1986) 37–60.
[9] A. Majumdar, B. Bhushan, Role of fractal geometry in roughness characterization and contact mechanics of surfaces, ASME J. Tribol. 112 (2) (1990) 205–216.
[10] A. Majumdar, C.L. Tien, Fractal characterization and simulation of rough surfaces, Wear 136 (2) (1990) 313–327.
[11] L. Kogut, R.L. Jackson, A comparison of contact modeling utilizing statistical and fractal approaches, ASME J. Tribol. 128 (1) (2005) 213–217.
[12] K.L. Johnson, J.A. Greenwood, J.G. Higginson, The contact of elastic regular wavy surfaces, Int. J. Mech. Sci. 27 (6) (1985) 383–396.
[13] J.F. Archard, Elastic deformation and the laws of friction, Proc. R. Soc. Lond. A 243 (1957) 190–205.
[14] A. Majumdar, B. Bhushan, Fractal model of elastic–plastic contact between rough surfaces, ASME J. Tribol. 113 (1) (1991) 1–11.
[15] E.J. Abbott, F.A. Firestone, Specifying surface quality—a method based on accurate measurement and comparison, Mech. Eng. 55 (1933) 569–572.
[16] F.M. Borodich, B.N.J. Persson, Comment on elastoplastic contact between randomly rough surfaces, Phys. Rev. Lett. 88 (6) (2002) 069601.
[17] M. Borri-Brunetto, A. Carpinteri, B. Chiaia, Scaling phenomena due to fractal contact in concrete and rock fractures, Int. J. Fract. 95 (1–4) (1999) 221–238.
[18] M. Borri-Brunetto, B. Chiaia, M. Ciavarella, Incipient sliding of rough surfaces in contact: a multiscale numerical analysis, Comput. Meth. Appl. Mech. Eng. 190 (46/47) (2001) 6053–6073.
[19] K. Komvopoulos, N. Ye, Three-demensional contact analysis of elastic–plastic layered media with fractal surface topographies, ASME J. Tribol. 123 (3) (2001) 632–640.
[20] T.L. Warren, D. Krajcinovic, Fractal models of elastic-perfectly plastic contact of rough surfaces based on the cantor set, Int. J. Solids Struct. 32 (19) (1995) 2907–2922.
[21] K. Willner, Elasto-plastic normal contact of three-dimensional fractal surfaces using halfspace theory, ASME J. Tribol. 126 (1) (2004) 28–33.
[22] W. Yan, K. Komvopoulos, Contact analysis of elastic–plastic fractal surfaces, J. Appl. Phys. 84 (7) (1998) 3617–3624.
[23] B.N.J. Persson, Elastoplastic contact between randomly rough surfaces, Phys. Rev. Lett. 87 (11) (2001) 116101.
[24] B.N.J. Persson, F. Bucher, B. Chiaia, Elastic contact between randomly rough surfaces: comparison of theory with numerical results, Phys. Rev. B 65 (2002) 184106–184111.
[25] M. Ciavarella, G. Demelio, J.R. Barber, Y.H. Jang, Linear elastic contact of the Weierstrass profile, Proc. R. Soc. Lond. A 456 (2000) 387–405.
[26] H.M. Westergaard, Bearing pressures and cracks, ASME J. Appl. Mech. 6 (1939) 49–53.
[27] M. Ciavarella, G. Murolo, G. Demelio, J.R. Barber. Elastic contact stiffness and contact resistance for the Weierstrass profile. 52 (6) (2004) 1247.
[28] L. Kogut, K. Komvopoulos, Analysis of spherical indentation cycle of elastic-perfectly plastic solids, J. Mater. Res. 19 (2004) 3641–3653.
[29] S.D. Mesarovic, N.A. Fleck, Frictionless indentation of dissimilar elastic–plastic spheres, Int. J. Solids Struct. 37 (2000) 7071–7091.

[30] N.A. Fleck, G.M. Muller, M.F. Ashby, J.W. Hutchinson, Strain gradient plasticity: theory and experiment, Acta Metall. Mater. 42 (2) (1994) 475–487.
[31] H. Gao, Y. Huang, W.D. Nix, J.W. Hutchinson, Mechanism-based strain gradient plasticity. I. Theory, J. Mech. Phys. Solids 47 (1999) 1239–1263.
[32] H. Gao, Y. Huang, W.D. Nix, J.W.J. Hutchinson, Mechanism-based strain gradient plasticity. II. Analysis, J. Mech. Phys. Solids 48 (2000) 99–128.
[33] I.A. Polonsky, L.M. Keer, Scale effects of elastic–plastic behavior of microscopic asperity contacts, ASME J. Tribol. 118 (1996) 335–340.
[34] R. Rodriquez, I. Gutierrez, Correlation between nanoindentation and tensile properties: Influence of indentation size effect, Mater. S. Eng. A 361 (2003) 377–384.
[35] J.G. Swadener, E.P. George, G.M. Pharr, The correlation of the indentation size effect measured with indenters of various shapes, J. Mech. Phys. Solids 50 (2002) 681–694.
[36] M. Zhao, W.S. Slaughter, M. Li, S.X. Mao, Material-length-scale-controlled nanoindentation size effects due to strain-gradient plasticity, Acta Mater. 51 (2003) 4461–4469.
[37] J.J. Quicksall, R.L. Jackson, I. Green, Elasto-plastic hemispherical contact models for various mechanical properties, IMechE J. Eng. Trib. J. 218 (4) (2004) 13–322.
[38] K.L. Johnson, Contact Mechanics, Cambridge University Press, 1985.