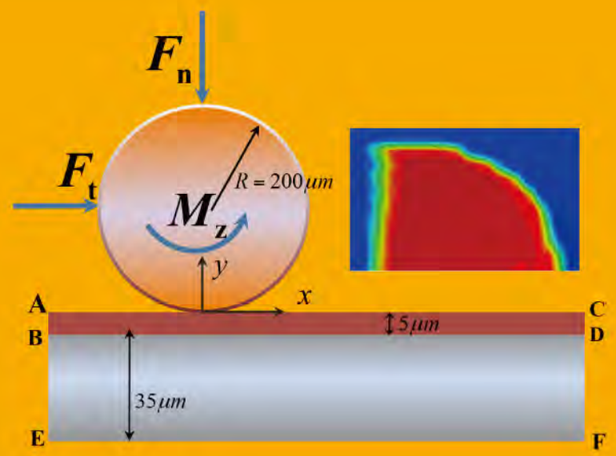
Friction

(Quarterly, Started in 2013)
Volume 5 Number 1 / March 2017

ISSN 2223-7690
CN 10-1237/TH



Tribology behavior on scratch tests: Effects of yield strength

摩擦（英文）（季刊，2013年创刊） 第5卷 第1期 2017年3月出版

Editor-in-Chief	Jianbin Luo	主管单位	教育部
Sponsored by	Tsinghua University	主办单位	清华大学
Supported by	Chinese Tribology Institute	学术支持	中国机械工程学会摩擦学分会
Edited by	Friction Editorial Office	主　　编	雒建斌
Published by	Tsinghua University Press	编　　辑	《摩擦》编辑部
Address	Xueyan Building, Tsinghua University, Beijing 100084, China	出版发行	清华大学出版社有限公司
		印刷单位	北京天成印务有限责任公司
Website	http://www.springer.com/40544 http://friction.tsinghuajournals.com		
Online Manuscript Submission, Review and Tracking System	http://mc03.manuscriptcentral.com/friction		

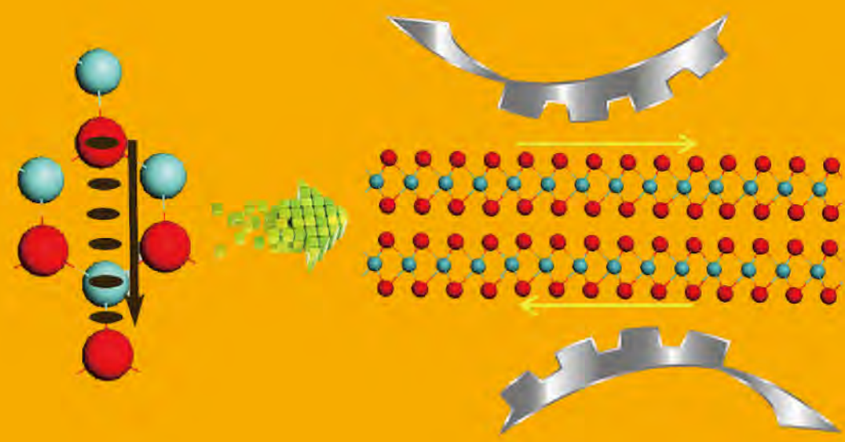
ISSN 2223-7690
03>
9 772223 769170

40544 Friction Volume 5 Number 1 2017 pp 1–122

ISSN 2223-7690
CN 10-1237/TH

Friction

Volume 5 Number 1 / March 2017



Computational investigation of the lubrication behaviors of dioxides and disulfides of molybdenum and tungsten in vacuum

TSINGHUA UNIVERSITY PRESS
Springer

Friction
ISSN 2223-7690
CN 10-1237/TH
Vol. 5, No. 1, March 2017

Contents

Review Article

Research Article

TSINGHUA UNIVERSITY PRESS Springer

Friction **5**(1): 1–22 (2017)
DOI 10.1007/s40544-017-0143-0
ISSN 2223-7690
CN 10-1237/TH
REVIEW ARTICLE

A review of friction models in interacting joints for durability design

Zulfiqar A. KHAN*, Vivek CHACKO, Hammad NAZIR

NanoCorr, Energy & Modelling (NCEM) Research, Department of Design & Engineering, Bournemouth University, Bournemouth, United Kingdom

Received: 31 July 2016 / Revised: 26 October 2016 / Accepted: 16 December 2016

Abstract: This paper presents a comprehensive review of friction modelling to provide an understanding of design for durability within interacting systems. Friction is a complex phenomenon and occurs at the interface of two components in relative motion. Over the last several decades, the effects of friction and its modelling techniques have been of significant interests in terms of industrial applications. There is however a need to develop a unified mathematical model for friction to inform design for durability within the context of varying operational conditions. Classical dynamic mechanisms model for the design of control systems has not incorporated friction phenomena due to non-linearity behaviour. Therefore, the tribological performance concurrently with the joint dynamics of a manipulator joint applied in hazardous environments needs to be fully analysed. Previously the dynamics and impact models used in mechanical joints with clearance have also been examined. The inclusion of reliability and durability during the design phase is very important for manipulators which are deployed in harsh environmental and operational conditions. The revolute joint is susceptible to failures such as in heavy manipulators these revolute joints can be represented by lubricated conformal sliding surfaces. The presence of pollutants such as debris and corrosive constituents has the potential to alter the contacting surfaces, would in turn affect the performance of revolute joints, and puts both reliability and durability of the systems at greater risks of failure. Key literature is identified and a review on the latest developments of the science of friction modelling is presented here. This review is based on a large volume of knowledge. Gaps in the relevant field have been identified to capitalise on for future developments. Therefore, this review will bring significant benefits to researchers, academics and industry professionals.

Keywords: friction; dynamics; joint clearance; numerical models; impact; durability

1 Introduction

Friction is a ubiquitous phenomenon which occurs at the interface of two surfaces in physical contact and in relative motion. It may be at times beneficial and/or detrimental in other scenarios. The phenomenon of friction is complex because it has time dependent non-linear characteristics and it is influenced by multiple factors. Friction phenomenon applies to scales ranging from nanometre level interactions to micron level interfaces to large geological interactions [1, 2]. Friction is directly linked to the durability and reliability of interacting systems and if it is not fully optimised then it leads to significant efficiency losses. According to the Jost report of 1966, "a sizeable portion of the GDP of a nation is spent in alleviating friction and its effects namely wear". Although tribology is a relatively new area, it is formed from a confluence of theory and empiricism, continued experimental analyses, mechanics, surface engineering, chemical interactions and more recently computational methodology. Since the phenomenon has both widespread and deep-rooted influence, this review paper seeks to gain an insight into the history of the development of friction and

* Corresponding author: Zulfiqar A. KHAN, E-mail: zkhan@bournemouth.ac.uk

List of symbols

A	Parameter in Bliman-Sorin model
B	Parameter in Bliman-Sorin model
$\boldsymbol{B}(\dot{\boldsymbol{\theta}})$	Frictional torque matrix
c_{d}	Dynamic friction coefficient in modified coulomb friction law
c_{f}	Friction coefficient in modified coulomb friction law
C	Parameter in Bliman-Sorin model
$\boldsymbol{C}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}})$	Coriolis and centripetal effects
C_i	Coordinate systems of the multibody with clearance
D	Damping coefficient
$\boldsymbol{D}(\boldsymbol{\theta})$	Inertia matrix
DLC	Diamond like coating
DV	Limit velocity in Karnopp model
E^*	Effective modulus of elasticity
e_y	The distance between centres along the ordinate
e_x	The distance between centres along the abscissa
e_{ij}	Vector distance along displacement of centres
F	Friction force generated by friction model
F_1	Force applied on the rigid body 1
F_2	Force applied on the rigid body 2
F_{c}	Coulomb friction force
F_{d}	Hysteresis force in the Leuven Model
F_{ext}	Force applied by external actuator
F_{f}	Friction force (Seven Parameter model)
F_{h}	Hysteresis force in the Leuven Model
F_{n}, F_{N}	Normal force at contact
F_{s}	Static friction coefficient or Stribeck friction coefficient
$F_{\mathrm{s,a}}$	Magnitude of the Stribeck friction at the end of the previous sliding period (Seven Parameter model)
$F_{\mathrm{s},\infty}$	Magnitude of the Stribeck friction after a long time at rest (Seven Parameter model)
F_{T}	Tangential friction force in modified coulomb model
F_{v}	Viscous friction force
$\boldsymbol{G}(\boldsymbol{\theta})$	Gravity forces
k	Spring constant
K	Stiffness at contact
$\boldsymbol{M}$	Mass matrix
n	Normal along the contact
n	Shape curve transitioning coefficient
N	Exponential coefficient
O	Origin of the global coordinate system
O_i	Coordinate system i of the journal
O_j	Coordinate system j of the journal
PEEK	Polyether ether ketone
$\ddot{\boldsymbol{q}}$	Generalised acceleration state vector
r	Clearance between the bodies at contact
R_{o}	Radius of the outer bearing
R_{i}	Radius of the journal
RI	Reset Integrator model
s	Space variable in Bliman Sorin model
$s(v)$	Shape transitioning curve
t	Tangent at contact point
t_2	Dwell time (Seven parameter model)
T_{L}	Time constant of frictional memory
v	Velocity at contact of the moving body
v_0, v_1	Threshold values of velocities for dynamic correction factor in the modified Coulomb friction model
v_{s}	State variable in Bliman Sorin model
v_{T}	Relative tangential velocity at contact
$\dot{x}$	Sliding velocity
$\dot{x}_{\mathrm{s}}$	Characteristic velocity of the Stribeck friction
x	Sliding distance
X	Hysteresis damping factor
X	X axis of the global coordinate system
X_1	Displacement of the rigid body 1
X_2	Displacement of the rigid body 2
Y	Y axis of the global coordinate system
z	State parameter in friction model
α	Baumgarte coefficients
β	Baumgarte coefficients
$\boldsymbol{\Gamma}$	Input matrix
γ	Time parameter of the rising static friction (Seven parameter model)
$\dot{\delta}$	Time derivative of deflection at contact
δ	Deflection at contact/Penetration depth of journal and bearing
λ	Lagrange multiplier
$\boldsymbol{\phi}_q$	Jacobian matrix for constraint equations
σ_q	Coefficient accompanying the state variable, an equivalent stiffness for position-force relationship at velocity reversal (LuGre model), the tangential stiffness of the static contact
σ_1	Micro-viscous friction coefficient
σ_2	Viscous friction coefficient
μ	Coefficient of friction
$\boldsymbol{\tau}$	Joint torque
$\boldsymbol{\tau}_{\mathrm{L}}$	Torque from external load

dynamic modelling and to summarize various friction models and their characteristics.

Friction occurs in both prismatic and revolute mechanisms contacts. In revolute joints, increasing the diameter of the contacts can effectively reduce the contact pressure. However, the sliding distance increases which may result in accelerated wear [3]. The nature of contact in revolute joints in manipulators can vary between conformal and non-conformal contacts depending on whether sliding bearings or anti-friction bearings have been used. The nature of the clearance existing at the revolute joint contact determines whether the contact is continuous or non-continuous contact during its operation. Continuous contacts can be modelled with a revolute friction model. The non-continuous models require contact models that capture the model dynamics as well as follow energy conservation and are therefore much more tedious to model. Both these models have been examined in Sections 4.1 and 4.2 of this paper.

The focus of several recent researches has been the modelling of friction in manipulators [4–8]. The extended problem also requires the formulation of a suitable control system. Some researchers have tried to use an un-modelled dynamics approach [9, 10]. Friction introduces non-linearity into the dynamics equation, which physically implies phenomena such as stick-slip in relative motion, limits cycles and introduces difficulties in positioning the end effector of the manipulator. However, as of now, friction models are imperative in analysing any mechanism.

Friction modelling has progressed from specific models analysing friction at the interface of geometries [11–14], to the analysis of friction at manipulator joints [4, 7, 15] with clearance and their kinematics and dynamic [16–25]. Marques et al. [26] have recently surveyed friction models in single degree of freedom in planar systems. Lately, researches such as Mukras et al. [27, 28] have analysed the computation of joint wear calculation along with dynamics. The progress of research is seen in the integration of multiple disciplines that include tribology, computational mechanics, control systems, surface interaction and chemical interactions.

This review converges to manipulator joints used in mechanical equipment as excavators and search and rescue smart mechanical systems. Excavator is a commonly deployed platform in disaster sites. However there has been a rise in accident numbers in controlled construction environments with respect to the number of units being used, which has been a major health and safety concern [29]. The manipulator kinematics has been introduced by Koivo et al. [30], extended to dynamics by Vähä and Koivo [31, 32]. Subsequent works have followed the modelling approach proposed by these researchers in attempting to develop the dynamics and control methods however until date only some studies based on the real arm [33, 34] have shown partial success in implementation. The non-linearity of the dynamics formulation makes the numerical solution both complex and computationally expensive. The computational effort increases with the increase in the degree of freedom, e.g., increase in the number of links in the manipulator, transformation from simple open chain manipulator to a closed loop mechanism and with the introduction of the nonlinear friction component into the dynamic equation. According to Haessig and Friedland, "friction is the nemesis of precision control". The phenomenon of friction is often ignored in control theory because of its intricacy. For precision control applications, however, the effect of friction cannot be ignored. The main impediment can be attributed to the complexity of dynamics, i.e., the non-linearity in the loads and more importantly the question of mimicking a human being. In this respect it is worth noting that Bilandi et al. [4, 15, 35] have studied the friction in an excavator arm.

Moreover, physical failure of a robot is a major obstacle in search and rescue missions [36, 37] and this can only be alleviated through the study of the manipulator mechanism design from the material science viewpoint, the need for which can be substantiated by the growing attention on natural and man-made disasters and the efforts to minimise causalities. The ingress and egress of rescuers is not the only cause for concern in such sites, and bringing such sites back to normalcy is part of the post-disaster operation. For these the use of equipment is very much a necessity both to speed up operation and to reduce risk to human beings however the risk reduction also entails focus on the manipulator mechanism to perform in those environments without catastrophic failures. Very few researches have focussed on this

aspect since the focus of search and rescue operations has been to detect and replace of live rescuers (human and dogs) with robots. When large quantities of chemicals were found in large radius after the WTC incident [38], the effect of corrosion inducing species on the operating equipment needs to be examined too. Stalwart researchers such as Blau [39] recognise that most appropriate method for determining the effect of friction and its effects and quantifying it, is still in experimentation and analysis also elucidates that the effects of environment on such mechanical joints need attention. Recently, holistic models combining dynamics, friction and wear have begun to appear in literature.

A multi-disciplinary approach (Fig. 1) is needed to fully analyse the problem and to devise a meaningful solution for the dynamics and control of manipulators. Friction and wear effects in the manipulators incorporating environmental effects need to be fully studied. With the increase in available computational power, a transition from simple analytical to complex numerical formulations of friction problem, with an analogous improvement in the range and precision of friction models has been looked. As part of the effort, a detailed literature survey is presented here, which provides an in-depth insight into the modelling of dynamics with a focus on friction (Fig. 2). A similar review for biodiesels has been presented recently [40]. The future research directions and gaps have been identified and presented for future reference.

Multidisciplinary techniques to advance design methods to improve efficiency, reliability and durability of contacting surfaces are explored in the literature [41–44]. The outline of the survey methodology has been given in the next section.

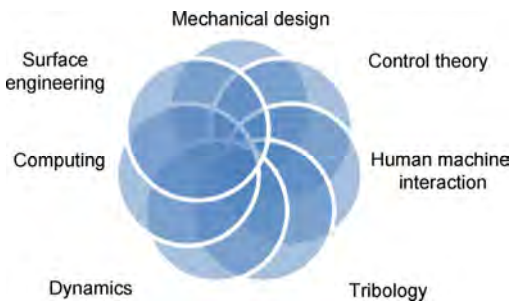


Fig. 1 The area of manipulator design modelling and control arises from the confluence of several branches of engineering and science.

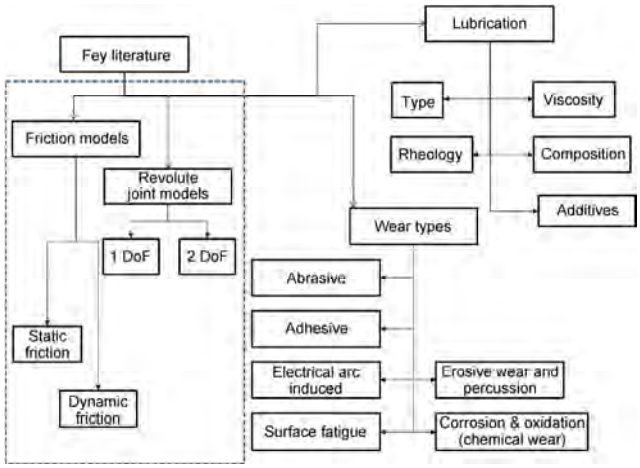


Fig. 2 Factors influencing the manipulator mechanism. This blue box highlights the focus areas of this paper, i.e., dynamics and friction with reference to manipulator methodology.

2 Review methodology

The review of literature was conducted beginning from important literature by citation index and relevance. Amstrong-Helouvry et al. [45] revealed that a growth of 700 articles in tribology is expected yearly. With the available volume of literature, it would be an impossible task to encompass all research areas. This review focuses on dynamics and friction modelling which applies to the specific case of the manipulator arm deployed in harsh environments. Kinematics, dynamics and control of such robotic manipulators have been the subject of research interest in the recent few decades. Important keywords are identified (refer to keywords above) in relevant cited publications. A search on (friction models*, static* and dynamics*), in June 2016 revealed the following statistics (Fig. 3). The focus of this paper is on the development of the techniques of modelling mechanisms, the recognition and the inclusion of friction into dynamic modelling, some necessary aspects of control and the evolution of the modelling methodology of friction along with wear and lubrication which form an integral portion of this science. At the outset, the following aspects are addressed, including (1) to summarize the research development and timeline, (2) to identify key review papers, (3) to enumerate the important numerical models and (4) to identify future research directions. The history of manipulator modelling is outlined in the forthcoming section.

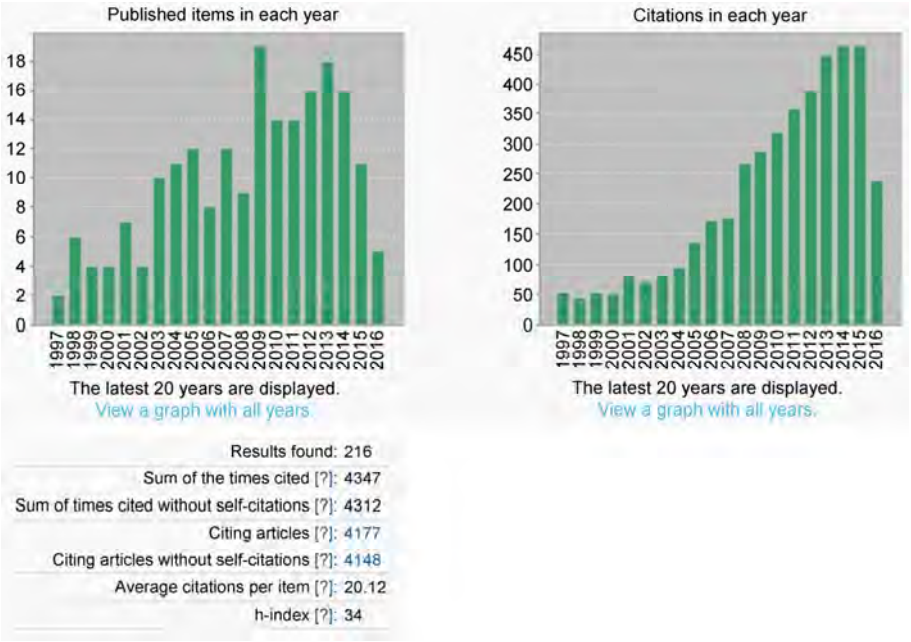


Fig. 3 Articles published in the domain containing keywords of friction models, static and dynamic.

3 History of dynamic modelling

An encapsulated version of the history of mechanics of manipulators and numerical modelling is presented here. Progress in manipulator modelling can be seen with the increase in modelling complexities from the late 1980s to present date. History of multibody dynamics has been presented by Rahnejat [46] and Schiehlen [47]. Computational dynamics has made rapid progress in the 20th century. Detailed modelling methodology from the robotics and control perspective has been presented by Siciliano and Khatib in their book [48]. Uicker et al. [49] have provided the fundamental theory of mechanisms. Computational aspects have also been described by Groover and Zimmers Jr. [50]. Dynamics of parallel manipulator with friction has been presented by Farhat et al. [51]. Friction in space manipulator has been presented in Hachkowski et al. [52]. With the advent of computers and increased availability of computing power, several techniques incorporating engineering design techniques have evolved. The use of CAD and multibody dynamics in the design, simulation and analysis of mechanisms has greatly contributed to the efficiency of the entire process [53].

The free body diagram of a robot with manipulator is given in Fig. 4. The first step in the modelling of any mechanism is the development of the kinematic relationship between links, assigning the appropriate relations between links. Planar kinematics of a manipulator arm with three revolute joints has been presented in Ref. [30] following the Denavit-Hartenberg [54] convention of coordinate system assignment. Computation of the forward kinematics of such mechanisms is straightforward. However, in the case of the inverse kinematics of multi-link mechanisms multiple solutions exist and computation of the inverse kinematics is difficult.

Kinematic analyses are devoid of force calculations. Dynamic analyses which include the force calculations are presented in literature [56, 57]. Dynamics of mechanisms can be modelled by using Newton Euler method [31], Euler Lagrange method [58], Gibbs Appel method [59], or Kane's equation [60, 61]. Selection of method used to model the system dynamics depends on the application and complexity of the mechanism design. The dynamic model of the manipulator is based on Ref. [31]:

$$D(\boldsymbol{\theta})\ddot{\boldsymbol{\theta}} + C(\boldsymbol{\theta},\dot{\boldsymbol{\theta}})\dot{\boldsymbol{\theta}} + G(\boldsymbol{\theta}) + B(\dot{\boldsymbol{\theta}}) = \Gamma\boldsymbol{\tau} - \boldsymbol{\tau}_{\mathrm{L}} \tag{1}$$

where $\boldsymbol{\theta} = [\theta_1\ \theta_2\ \theta_3\ \theta_4]^{\mathrm{T}}$ is the vector representation of joint angles, $D(\boldsymbol{\theta})$ represents inertia, $C(\boldsymbol{\theta},\dot{\boldsymbol{\theta}})$ represents Corioli's and centripetal effects, $G(\boldsymbol{\theta})$ represents gravity forces, $B(\dot{\boldsymbol{\theta}})$ represents frictional

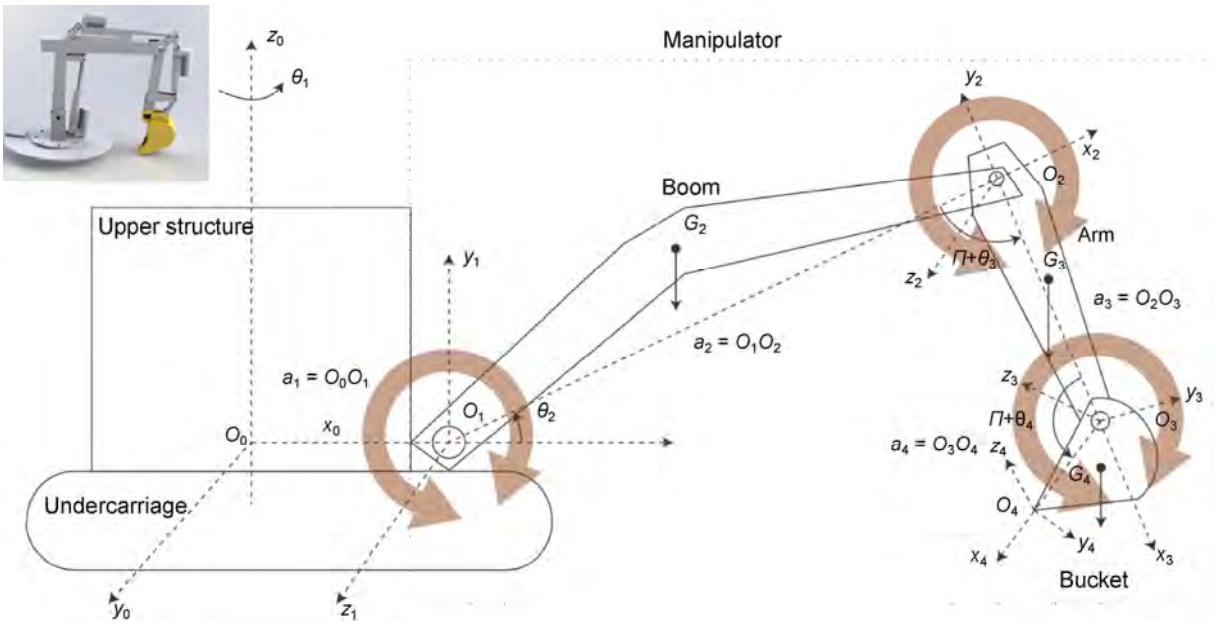


Fig. 4 The layout of a robot with a planar manipulator mechanism based on Ref. [55], the circle with arrows depicts the friction torques at rotary joints. The inset of the figure shows the rendering of a manipulator model generated within the computer aided design environment.

forces, $\boldsymbol{\Gamma}$ is the input matrix corresponding to joint torques $\boldsymbol{\tau}=[\tau_1, \tau_2, \tau_3, \tau_4]^{\mathrm{T}}$ and $\boldsymbol{\tau}_{\mathrm{L}}$ represents equivalent soil-tool interaction torques. The first angle, $\boldsymbol{\theta}_1$ represents the rotation of the manipulator about the base of the excavator, which is usually assumed to be null magnitude since the manipulator operation is assumed to be immobile in that degree of freedom. This means that the manipulator remains planar during digging since it does not turn about the base during this task. Therefore, the model complexity and the computational effort are reduced.

The soil tool interaction force F_{L} is a highly non-linear component, which acts on the end effector. Several researches are dedicated to the computation of soil-tool interaction forces [62–65]. The influence of the soil-tool interactions on the state variables of the manipulator would also affect the friction torque generated at the revolute joints. Therefore, simple friction models would be insufficient to capture the resulting frictional dynamics. In joint mechanism friction forces may be as high as 20% of the actuation force [7]. Simplification schemes may include (1) simplifying dynamics by ignoring some terms and correcting errors using feedback (e.g., non-linear friction effects, Coriolis's force and centripetal force which can be ignored at low link velocities but constitute a considerable component of forces at high speeds. The Coriolis's/Centripetal components cannot be corrected by feedback method.), or (2) tabulation lookup. Tabulation and interpolation method can be used to create a lookup table for pre-calculated values. Therefore, this technique cannot be employed when non-linear terms occur. Tabulation method cannot be applied to friction due to its high non-linearity. Recursive Newton formulation is more efficient than recursive Lagrangian formulation. However, they can be brought down to approximately the same computational time, making real time solution possible [66]. A diagrammatic representation of dynamics and control is given below in Fig. 5.

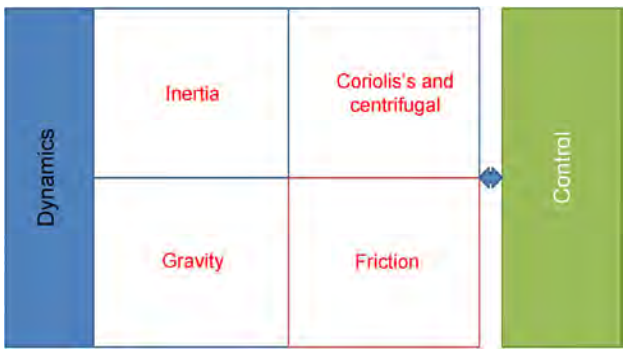


Fig. 5 Representation of the dynamics equation and its components.

For dynamic systems with clearance, the combination of differential and algebraic equations (DAE) resulting in the equation of motion is given by [67]:

$$\begin{bmatrix} \boldsymbol{M} & \boldsymbol{\phi}_{\mathrm{q}}^{\mathrm{T}} \\ \boldsymbol{\phi}_{\mathrm{q}} & 0 \end{bmatrix} \begin{bmatrix} \ddot{\boldsymbol{q}} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{g} \\ \boldsymbol{\gamma} \end{bmatrix} \tag{2}$$

where $\boldsymbol{M}$ is the mass matrix, and $\boldsymbol{\phi}_{\mathrm{q}}$ is the Jacobian matrix for the constraint equations. $\ddot{\boldsymbol{q}}$ includes the generalised state accelerations. $\boldsymbol{\lambda}$ denotes the Lagrange multipliers, $\boldsymbol{g}$ is the generalised force vector, and $\boldsymbol{\gamma}$ represents quadratic velocity terms dependent on velocity, position and time. This equation can be solved by using solution methods which are applicable to algebraic equations in the absence of redundant constraints. Baumgarte method is the most widely folloed stabilisation methods in solution of DAEs. In the case of redundant constrains, the augmented Lagrangian method is employed. Detailed formulation method is presented by Flores et al. in Ref. [67] and the model is employed in the majority of subsequent modelling works.

In the next section the progress in friction modelling has been presented.

4 History of friction modelling

While modelling the spatial behaviour, and the dynamics of mechanisms has made significant progress, it is necessary to capture the effects of friction in the joints. In classical modelling, the effect of friction is not considered. However, friction is defined as the tangential reaction force that occurs between two surfaces in contact, dependent on factors that include the contact geometry, the topology, relative velocity of surfaces in contact and displacement of surfaces, load and lubrication [23, 68–71]. Friction is a complex phenomenon caused by the interaction of the surface and near surface regions of two interacting components as well as lubricants if present between such surfaces [72]. However, the classical friction model does little more than to give an approximation of friction forces in static analyses.

The selection of friction models is based on the operational condition during application. Several mentioned models include evaluation of physical friction and wear. Olson et al. [68] have examined several friction models which are available in the context of automatic control. Within the domain of control theory friction effects are addressed in dynamics by quantifying parameters as noise generated from the ensuing effects [73], however such an approach falls short to address the overall phenomena of friction and wear from the mechanical design approach.

In the forthcoming section, static and dynamic friction models are examined, with their brief history and modelling equations.

4.1 Static friction models

Literatures [2, 74] reveal preliminary inquiries into the nature of friction of interacting bodies. The postulates of friction according to Guillaume Amontons [75] are given as:

- The force of friction is directly proportional to the applied load, i.e., $\mu \propto W$.
- The force of friction is independent of the apparent area of contact.
- Kinetic friction, μ_{k} is not proportional to (independent of) the sliding velocity.

Therefore, the simplest representation of friction can be given as

$$F_{\mathrm{f}} = \mu W \tag{3}$$

where the coefficient of friction μ, is dependent on the mating materials in interaction, surface preparation and operating conditions. The force required to initiate movement is known as the static friction force. The force required to maintaining motion is called kinetic friction force. These different magnitudes with the value of frictional force at limiting conditions have a greater value compared to kinetic condition.

Coulomb proposed the simple roughness model in 1785 which is used for friction force calculation. It is a static model that has neither history nor states [7] and may be explained based on the quasi-static properties of materials. The Coulomb model has no dissipative component to it, which is a drawback [76]. It is given by Ref. [77].

$$F_{\mathrm{c}} = \mu F_{\mathrm{n}} \mathrm{sign}(v) \tag{4}$$

Also, shear failure is the predominant cause in sliding

with friction. For static and Coulomb friction, the friction forces are proportional to the normal load. At low velocities, the shear strength of a solid lubricant film is high compared to the corresponding shear forces of the fluid film building up at higher velocities. The viscous friction can be represented by

$$F_v(v) = \sigma_v v \tag{5}$$

If the lubricating film is sufficient to separate the bodies in contact completely, the hydrodynamic effects become significant, i.e., the friction coefficient may increase with the velocity. Therefore, the friction force generated in lubricated systems normally decreases when the velocity increases from zero. This is called the Stribeck effect.

While the Coulomb and viscous friction models account for the fundamental modelling of friction at joints, the addition of Stribeck friction accounts for low velocity, high magnitude friction. The combined effects of the Coulomb, viscous and Stribeck components of friction is shown in Fig. 6. The mathematical representation of the combined effects of static friction models is given by:

$$F_f(v) = \mu F_n \text{sign}(v) + \sigma_v v + F_s(v) \tag{6}$$

These effects, which have been evaluated initially for the linear sliding models, also apply to revolute models as revolute friction torque. Above mentioned model does not accurately capture friction and its effects. The requirement of dynamic friction models is highlighted which have been explained in the next section.

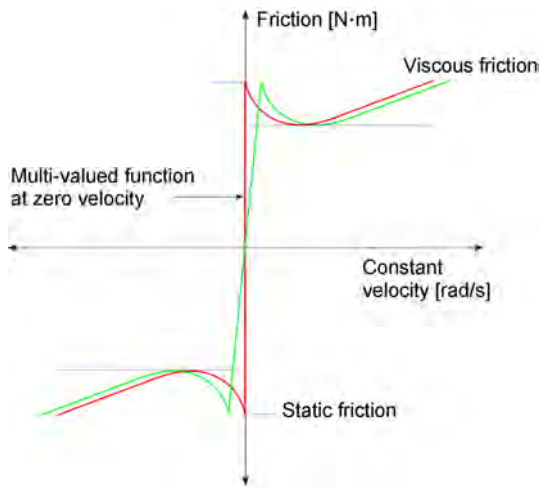


Fig. 6 The combined effect of Coulomb friction, viscous friction and low velocity Stribeck effect based on Ref. [78].

4.2 Dynamic friction models

The phenomenon of friction is being increasingly applied to modelling of dynamic systems and their control. Several models in literature include Dahl model (1968), Karnopp (1985), Bliman Sorin (1995), LuGre (1998) and the Leuven (2000) model. Classical friction models do not accurately predict limit cycle. For precision control applications, however the effect of friction cannot be ignored, i.e., micron level motion gradient in manipulator junctions may result in positioning various several orders higher at the end effector location. Physics motivated models such as generalised Maxwell slip model, Frenkel Kontrova, Tomlinson, Frenkel Kontrova Tomlinson model, Barridge Knopoff model, and Tomlinson models [77] are not presented here. These models improve accuracy of modelling however cannot be employed for control systems owing to their computational time requirement. Several friction models have been presented in literature incorporating dynamic friction models to enhance capturing the effects of friction. Friction model must account for the transition phenomena between static and kinetic contact, and account for hysteresis effect and direction reversals. These models have been introduced to bridge the gap in performance found in the static friction models, which have been explained in the previous section.

Developments in the dynamic friction models have occurred perhaps from the mid half of the twentieth century. The important models have been summarised in Eqs. (7)–(12). The key aspects included in these models are the genesis of friction, stick-slip phenomenon, hysteresis, friction lag and friction memory which make it more complicated and at the same time capture friction effects better compared to static models which are presented in Section 4.1. Dahl model is among the first of such models [79]. According to Olsson et al. [68] both the LuGre model and the Bliman-Sorin developments on the Dahl model capture viscous friction, Stribeck friction and the phenomenon of stiction. The LuGre model adds the effect of damping to the Dahl model. Piatkowski [80] provides a recent analysis of Dahl and LuGre models. Dynamic

friction models have been considered by Karnopp [81], Quinn, Kikuuwe [82, 83], Awrejcewicz and Kudra [84], and Wojewoda et al. [85].

Dahl Model-1968 [79, 80]:

$$F = \sigma_0 z \tag{7}$$

$$\text{where } z = v\,\text{sgn}\left(1-\text{sgn}(v)\frac{\sigma_0 z}{F_c}\right)\left|1-\text{sgn}(v)\frac{\sigma_0 z}{F_c}\right|^{\delta_p}$$

Karnopp Friction model-1985, 2008 [81, 86]:

$$F_{\text{friction}} = \begin{cases} -F_c\cdot\text{sgn}(\dot{x}) - F_v\dot{x} & |\dot{x}| \geqslant \text{DV} \\ -(F_{\text{ext}} - kx) & \begin{cases} |\dot{x}| < DV \\ |F_{\text{ext}} - kx| \leqslant F_s \end{cases} \\ -F_s\cdot\text{sgn}(F_{\text{ext}} - kx) & \begin{cases} |\dot{x}| < DV \\ |F_{\text{ext}} - kx| > F_s \end{cases} \end{cases} \tag{8}$$

where F_c is the Coulomb friction coefficient, F_v is the viscous friction coefficient, F_s is the static friction coefficient, DV is the limit velocity and the sgn(.) function is given by

$$\text{sgn}(\alpha) = \begin{cases} 1 & \alpha > 1 \\ 0 & \alpha = 1 \\ -1 & \alpha < 1 \end{cases}$$

Bliman-Sorin Model-1995 [87, 88]:

$$\frac{dx}{ds} = Ax_s + Bv_s \tag{9}$$

where $F = Cx_s$

$$A = \begin{pmatrix} \frac{-1}{\eta\varepsilon_f} & 0 \\ 0 & \frac{-1}{\varepsilon_f} \end{pmatrix},\quad B = \begin{pmatrix} f_1/(\eta\varepsilon_f) \\ -f_2/\ \varepsilon_f \end{pmatrix},\quad C = (1\ \ 1)$$

LuGre Model-1998 [68, 89]:

$$F = \sigma_0 z + \sigma_1\dot{z} + \sigma_2\dot{z} \tag{10}$$

where $\dot{z} = v - \frac{\sigma_0 |v|}{s(v)}z$

$$s(v) = F_c + (F_s - F_c)\exp((-v/v_s)^{\delta_{vs}})$$

Leuven Model-2000 [72, 90]:

$$\frac{dz}{dt} = v\left(1-\text{sgn}\left(\frac{F_d(z)}{s(v)-F_b}\right)\left|\frac{F_d(z)}{s(v)-F_b}\right|^n\right) \tag{11}$$

$$F_f = F_h(z) + \sigma_1\frac{dz}{dt} + \sigma_2 v$$

$F_h(z)$ is the hysteresis force, n is a coefficient used to transitions the shape curves, $s(v)$ is function which models the constant velocity behaviour, given by:

$$s(v) = \text{sgn}(v)(F_c + (F_a - F_e)e^{-(|v|/V_a)^\delta}$$

Seven Parameter Friction model 1994 [45, 77]:

$$F_f(x) = -k_t x \ \text{(presliding)} \tag{12}$$

$$F_f(v(t),t) = -(F_c + F_v\,|\,v(t)\,|) - \text{sgn}(v(t)F_s(v,t_2)\frac{1}{1+\left[v\frac{(t-\tau)}{v_s}\right]}$$

(Coulomb and viscous sliding)

$F_{s,a_n} = F_{s,a_{n-1}} + (F_{s,\infty} - F_{s,a_{n-1}})\frac{t_2}{t_2+v}$ (Rising static friction-breakaway)

Piedboeuf et al. [7] proposed an algorithm for computing joint friction in robotic simulations which includes Stribeck regime along with the stick-slip process validated against a planar robotic arm. At zero velocity, the friction value may be any value between $\pm F_s$ as shown in Fig. 6. To alleviate the problem introduced by bi-valued function at zero velocity, a gradient is introduced between the transition [91] and is reflected in Fig. 6. The solution to overcoming this non-unicity is to insert a linear slope across zero crossing where the function becomes bi-valued (applied similarly in [91, 92]).

Dahl Model is similar to the reset integrator model and includes (1) mechanism for zero velocity sticking and (2) application independent design. Disadvantages of the model are that (1) it does not generate a stiction force exceeding the sliding force unlike the other 4 models but can be modified to accommodate it and (2) lower accuracy compared to Karnopp and RI model.

Karnopp used bond-graph method to model the effects of friction for a two body system [81]. The model can be represented by Eq. (8). While this captures the energy aspect of the system, the disadvantage of the bond-graph equation is that it has to be formulated for

every model. The order of Karnopp model [81, 86, 93] reduces at zero relative velocity between the surfaces. Advantages of this model include (1) stick-slip phenomenon included in the model and (2) 30% faster execution. The drawbacks of the Karnopp model are that (1) the complexity of the model increases with increasing complexity of the dynamic system and (2) all combination possibilities of motion between the bodies must be considered. More detailed friction models such as the Dahl model and LuGre model (Eq. (10)) which account for the pre-sliding conditions have been presented in literature. Both Dahl and LuGre model are rate dependent because of which they cannot capture the reversal point memory. Swevers et al. [72] uses the LuGre model which performs satisfactorily for constant sliding velocity and suggests modification to it. In LuGre model the parameter z can be interpreted as the average bristle deflection. A change in the magnitude of frictional force occurs due to the transition between static and kinetic phase breakaway phenomenon occurring. Therefore, the transitional friction needs to be considered. However, LuGre model does not account for hysteresis behaviour. Swevers et al. [72] model includes Stribeck friction in sliding, hysteretic behaviour in pre-sliding, frictional lag, varying breakaway and stick-slip behaviour, supported by experiments but does not account for material characteristics and the effect of loads on material variations. The Bliman Sorin model presented in Eq. (9) is modified form of LuGre model [77]. Leuven model Eq. (11) is a modification of LuGre model and includes hysteresis with nonlocal memory. The modified Leuven model presented in Ref. [90] addresses the issues of memory stack size and frictional force discontinuity at closure of frictional loop with Lueven model. De Wit et al. [94] described the loss of performance of high precision manipulators owing to the effects of friction. The effect of friction lag and the existence of a hysteretic relationship between friction and velocity are considered. Breakaway force can also vary according to dynamics of the contact. At microscopic contact dimensions the velocity between the contacts will be non-zero. Stick slip motion is also seen in joints. Friction compensation through observer method is used to develop the control system in such a case. The model captures friction phenomena while maintaining simplicity. The performance of the contact through start of motion to its end and the performance at various velocities have been incorporated into the system.

Haessig and Friedland [95] present two friction models of which one is based on the bristle formulation intended to capture "sticking" effect (Fig. 7) and the other is called "reset integrator" model which does not encompass sticking but is similar to Karnopp model. The process of initiation of friction is described as the interaction of peaks that initially resembles a spring damper with high stiffness that is reluctant to allow motion. Discontinuity of friction at zero velocity causes very short computational time steps and steep slopes where a linear bypass is implemented.

Reset integrator (RI) model uses an auxiliary integrator to represent the phenomena of stiction. Advantages of this model are: (1) it is application independent; (2) it does not require re-derivation to suit each application; (3) it accurately represents bonding effect of stick-slip friction force; (4) this is a logical model; (5) loads are calculated to accommodate sticking loads and damping term; (6) damping mode is different from bristle in that there are no two separate modes; (7) it is computationally efficient; (8) short time steps due to breaking bristles are avoided; and (9) selection of parameters is much simpler than that for bristle model. The procedure is listed in Haessig and Friedland [95]. Seven-parameter model Eq. (12) consists of a spring model to capture pre-sliding and Coulomb, Viscous, Stribeck friction and friction lag.

The Bristle model presented in Fig. 7 is a simple algorithm which is (1) more efficient and accurate; (2) friction represented as many bristles which deflect with stiffness and damping, representative of surface

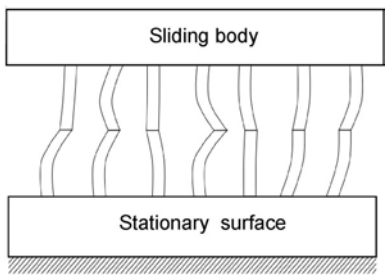


Fig. 7 The Bristle model shown above is one in which the contacting surfaces interact through bristles [95].

contact at joint; (3) frictional force is a function of velocity; (4) accurate model; (5) number of bristles control the fineness of the model. Disadvantages of the model are that (1) it is not efficient in terms of computational time; (2) fine spaced bristles cause successive short computational time steps and can decrease efficiency of the solution method or cause algorithm execution to fail; and (3) frictional force can become noisy signals.

Efficiency and accuracy of the models have been compared using fourth order Runge-Kutta method. The order of models in terms of computational efficiency are Dahl followed by Classical friction model, Karnopp, reset integral model and the bristle model. The selection of model is a trade-off between accuracy and computational efficiency, and the need for further comparison between RI and Dahl models is highlighted. The next section examines the issue of joint clearances in revolute joint contacts.

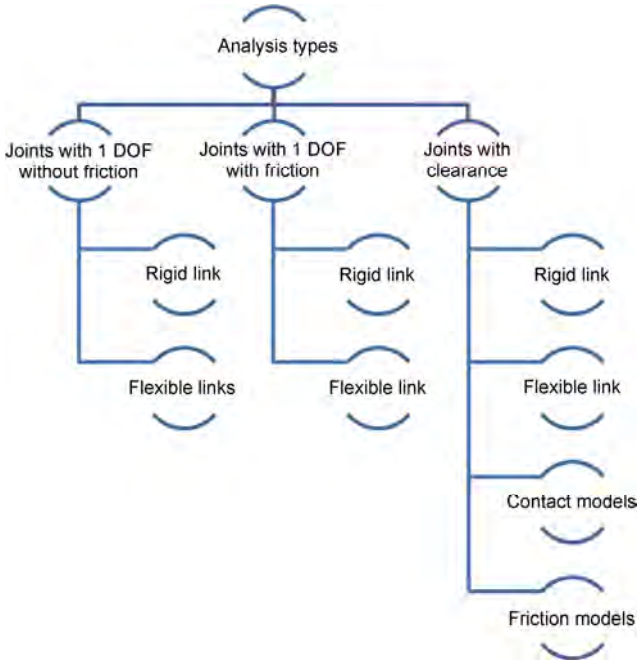


Fig. 8 Classification of manipulator dynamics research and problems.

5 Mechanistic models with clearance and friction

A classification of mechanism models has been presented in Fig. 8. While the geometric and kinematic analyses provide partial insight into the system performance, dynamic models are required to fully describe complex interacting systems. Joints are introduced to provide some constraint on the motion of the mechanism. The joints can be dry or lubricated and contact can be intermittent or continuous and is usually determined by the area of application. Major simulation studies in the area have employed dry friction model. In an ideal mechanism, the joints would have a close fit leading to the classical mechanism model. In actual joints, however the problem of joint clearances exists. Clearances in mechanical systems may occur due to assemblage and manufacturing errors, usually result in undue wear, performance loss, reduction in stability, noise, and dynamic impact loads, and affect the transfer of system loads [96–98]. This alters the performance of the mechanism and affects the dynamics, control and durability of the mechanism. One example of this is the problem of manipulator end effector positioning [98]. Several factors such as the contact stiffness, surfaces condition, and lubricant need to be considered while developing the contact model for a joint with clearance. In literature the revolute joints are considered either in stand-alone configuration [98] or in assembled form of the slider crank mechanism [99] and four bar mechanism [20]. Wang also addressed a similar problem from the perspective of robotic manipulator environmental interaction. Multiple friction contacts in mechanical systems have been analysed in Ref. [100].

Modelling and simulation of multibody dynamics with joint clearances is relatively new area of research [96]. Three reviews are identified from literature namely Haines [101]—unlubricated revolute joints (1979), Flores (2010) [21] and Machado et al.—compliant contact force models (2012) [102]. The problem of multibody impact with friction was first analysed by Routh in 1891 [100]. The model of dynamic systems with mechanical clearance presented by Dubowsky and Freudenstein [96, 103] in a two part publication and introduced the concept of impact pair model in which the surfaces in contact are modelled as compliant, i.e., as spring damper contacts as shown in Fig. 9. In 1975, Hunt and Crossley [104] studied the influence of the coefficient of restitution between two impacting bodies based on the force-law approach and recorded the results from the numerical simulations. The issue of impacting multi-bodies with kinematic

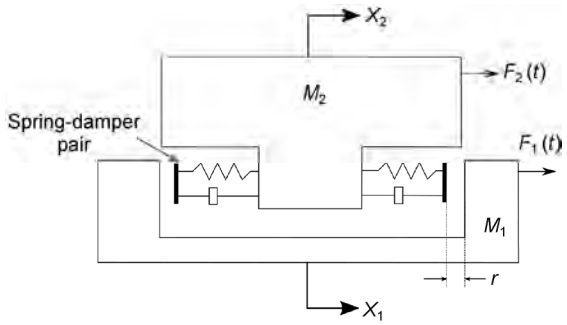


Fig. 9 Simple dynamic link coupling with clearance which leads to the "impact pair" condition based on Ref. [96].

contacts under the action of impulsive forces has been addressed by Lankarani and Nikravesh in [105]. Lankarani-Nikravesh contact model has been subsequently used in literature (Fig. 10). Rhee and Akay [106] investigated the revolute joint with clearance for a four bar mechanism and found a non-linear dependence on both the size of the clearance and coefficient of friction with a simple friction model for sliding friction. Friction and impact with joints clearance have been presented by Periera and Nikravesh [107] for intermittent motion with dry friction at the contact.

Spring-damper pair can be used to define connection between contacting surfaces, and are activated at the beginning of the contact. Contacting surfaces are initially assumed rigid. Energy dissipation cannot be modelled if the interaction is modelled exclusively by using only a spring connection because the spring has no dissipative component. To analyse the contact-impact the contact should be split into separate contact and departure events where each event is represented by switching function elements as shown in Fig. 11.

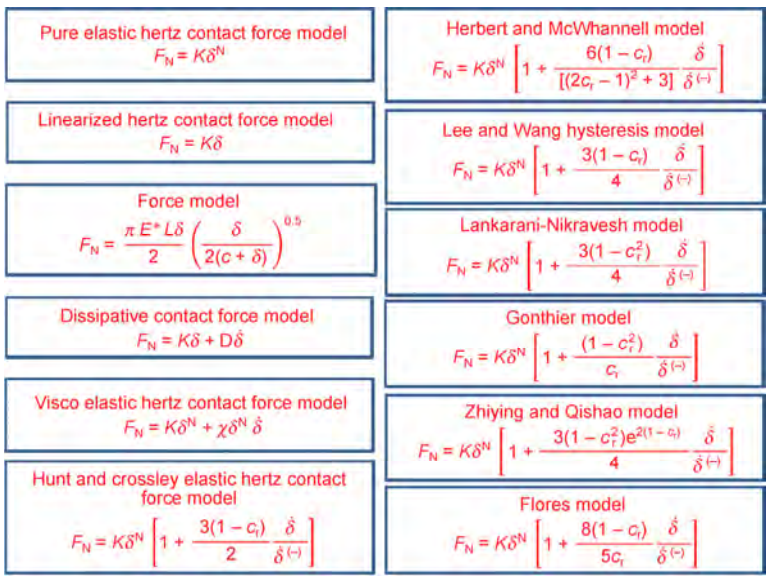


Fig. 10 Equations for the different contact models [102, 104, 108].

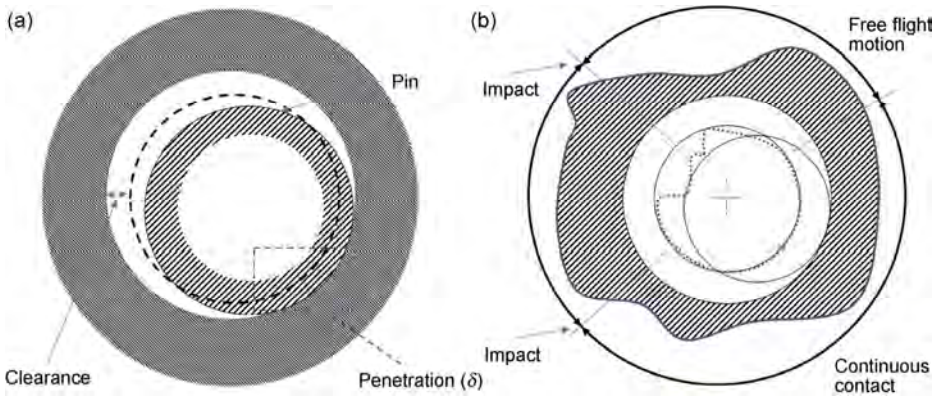


Fig. 11 The different modes in a revolute joint with clearance [109], and penetration in joints [110].

An alternative method of impact modelling for such contacts is the energy or Euler-Lagrangian method. Flores et al. [109] (see Fig. 12) has used momentum as a state variable of integration for two link pendulum and slider crank systems. Bauchau and Rodriguez [97] present a similar case in which finite element method (FEM) along with dynamic model has been proposed for a slider crank mechanism with flexible links.

Flores et al. [109] have analysed the dynamics of a slider-crank mechanism with clearance in the revolute joints. Hertzian contact model and Lankarani-Nikravesh contact impact model are used for calculating the contact parameters [111]. Koshy et al. [25] have evaluated the revolute joint with clearance for a rigid link slider crank mechanism focussing on the Hertzian contact model and extending the model to include damping and compared the results with experimental values. Hertz law for contacts is a static model as shown in Fig. 10, which has only a spring component that prevents energy dissipation. This violates the Law of Conservation of Energy at the contact. Hertz contact law is a nonlinear model. Energy transfer and conservation is a highly complex process. Lankarani-Nikravesh model is widely applicable compared to pure elastic force law models [25]. Acceleration parameters from the numerical simulations have been compared with values recorded from the slider crank mechanism. Clearance joints have an impact on the performance of mechanism as seen from experimental data [112]. It is worth noting that future analyses need to include joints material properties. Flores et al. [99] have concluded that lubrication alone alleviates much of the effects introduced by joint clearance owing to the inherent damping qualities. This implies that the lubricant properties influence the dynamics of the system. Zhao et al. [113] presents revolute joint's dynamics with mixed lubrication model by using Lagrange method and incorporating finite element method (FEM) for modelling small end of a connecting rod in an internal combustion engine. Machado et al. [108] have compared the performance of various contact force models graphically. Mukras et al. [27] has presented a combined model including joint elasticity and viscosity for mechanism dynamics and is shown in Fig. 13.

Other models in literature include massless link in 4 bar linkage with clearance [114], three step contact model with three configurations: (1) free flight, (2) impact, and (3) sliding. [115] presents the three step model of the slider crank mechanism with reaction only on contact. [116] presents a four-link mechanism with three step model, using discontinuous method for pre-collision and post collision momentum balance, three mode approach in which the impact and sliding computed by using a contact force model. Non-ideal joints use force constraints and are modelled by Ravn [110]. According to contact impact pair, flexible mechanisms with multiple clearance [117], joint clearance for massless link and clearance joints have been proposed by Earles and Wu [118]. Slider crank mechanism with multiple clearance joints has been modelled by Yaqubi et al. [119]. Flores et al. [99] has studied the performance of lubricated journal bearings and slider crank

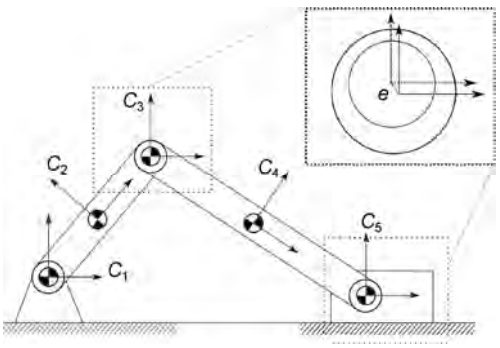


Fig. 12 Multibody system using slider crank mechanism with clearance.

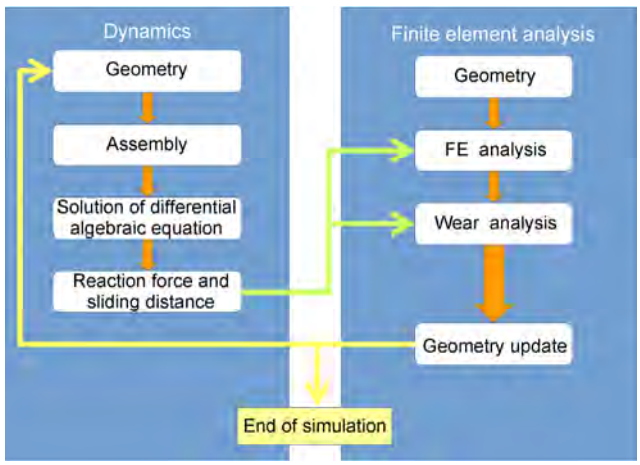


Fig. 13 Wear analysis integrated with the dynamics based on Ref. [27].

mechanism by using Pinkus-Sternlicht revolute model. Results from the publication shows that the model operates within the bounds of hydrodynamic lubrication theory. Elasto-hydrodynamic lubrication has been considered by Flores et al. [67] and Li et al. [120]. Several contact force models including pure elastic Hertzian contact force model, linearised Hertzian contact force model, force model, dissipative Spring Damper model, Gonthier model, Zhiying and Qishao model, Flores model, Visco-elastic Hertzian contact model, Hunt and Crossley model, Lee Wang hysteresis model, Lankarani-Nikravesh model and Hybrid model are presented in Fig. 10. A summary of these models and their successive improvements are presented in a survey of literature [98].

Three challenges that exist in multibody mechanical systems are (1) selection of appropriate constitutive law for the contact-impact event, (2) selection of appropriate contact stiffness and damping coefficient, and (3) quantification of energy transfer that occurs in such an event which leads to hysteresis [108]. Dynamics of collision may be classified as non-smooth dynamics formulation and the regularised approach (see Fig. 14).

A solution method for linear complementarity problem which can violate energy conservation principles has been used in Ref. [121]. Other solution methods include differential variation inequality (DVI) and Moreau's time stepping algorithm [21, 121]. However the limitation in Ref. [121] can be overcome

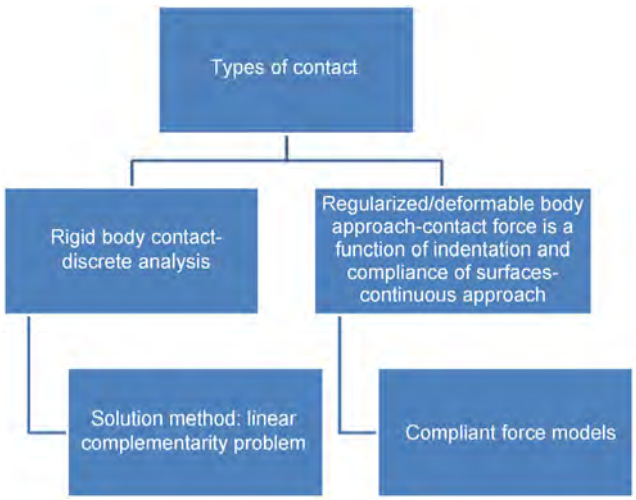


Fig. 14 Contact classification according to Machado et al. [102].

by choosing appropriate friction and degree of non-linearity for complex contact conditions. In addition, the problem of energy dissipation without violating the energy conservation condition is critical since even for low energy impacts, energy is transformed into sound and mechanical vibrations, as a function of the coefficient of restitution whose definition is subjective. The coefficient of restitution, which is representative of energy dissipation, is dependent on factors such as geometry of the contacting surfaces, pre-impact velocity, local material properties, duration of contact, temperature and friction [102]. Machado et al. [102] and Flores et al. [99, 102] have highlighted the importance of choosing the appropriate model for the mechanism. The Gonthier model [26] and Flores models are suitable for moderate to low coefficients of restitution. These modified models show enhanced accuracy and the possibility of a unified model requires further investigation.

The best solution technique can be chosen such that the simulation results can be validated. There is sufficient evidence of consistent improvement in available models, which bring simulations closer to reality.

In joints with clearance, the condition of impact with motion exists and the sum of forces at the instance of impact can be represented by:

$$F = F_{\mathrm{N}} + F_{\mathrm{T}} \tag{13}$$

Although several impact models are presented in literature to compute the normal force at the point of impact in Fig. 10, most literature utilises modified Coulomb friction law for computing tangential force at the point of impact. The predominant model for friction employed in dynamic models with clearance in joints is Coulomb dry friction model or modified Coulomb dry friction model [19, 21, 67, 122]. This can be explained by the fact that friction effect is only a minor component of the dynamic contact phenomenon in a joint with clearance and therefore is relegated in the analysis. Modified Coulomb friction law is given by [18]

$$F_{\mathrm{T}} = -c_{\mathrm{f}} c_{\mathrm{d}} F_{\mathrm{N}} \frac{v_{\mathrm{T}}}{\| v_{\mathrm{T}} \|} \tag{14}$$

where F_{T} is the tangential friction force, c_{f} is the friction

coefficient, c_d is the dynamic friction coefficient, F_N is the normal force, and v_T is the relative tangential velocity. Dynamic correction coefficient c_d is given by

$$c_d = \begin{cases} 0 & \text{if } v_T \leqslant v_0 \\ \dfrac{v_T - v_0}{v_1 - v_0} & \text{if } v_0 \leqslant v_T \leqslant v_1 \\ 1 & \text{if } v_T \leqslant v_1 \end{cases} \quad (15)$$

where v_0 and v_1 are the transition velocities.

The use of dynamic correction factor c_d improves time stepping characteristics of the solution algorithm. The influence of F_N on F_T is therefore determined by impact model. The analysis of journal bearing with clearance joint is presented by Bai and Zhao [98] which incorporates a new contact force model. The essence of applying any such model is to capture the actual physical phenomenon of impact, rebound and movement as shown in Figs. 11 and 15. Although several models have been reported in literature, there is scope for improvement in the numerical prediction of joint performances and this requires a case by case evaluation depending on the number of factors affecting contact conditions and performance. Since the contacting surfaces are influenced by several factors, the interacting surfaces and their durability have been examined in the following section.

6 Durability in harsh environments

Durability is the capacity of the mechanism to perform the designated function and fulfil the intended design life without unexpected failure. Durability is critical in all cases especially in mechanisms that are designed for deployment in harsh environmental and operational conditions. Failure of equipment leads to halting progress, loss in revenues and could cause accidents, e.g., search and rescue missions in disaster stricken areas. Therefore design life cycle analysis is critical for specialised deployment in high risk environments [123].

The failure of mechanisms deployed in hazardous environments is discussed in literatures [36, 124, 125]. Design failure was recognised as a major factor. This also includes the failure of components. Mechanical durability of mechanisms subject to both constant and variable loading is highly desirable. Interacting surfaces of the manipulator revolute joint between various links are affected by multiple factors such as load, sliding speed, lubrication, heat and the influence of external agents such as corrosive fluids and debris. Therefore, failure in interacting surfaces may occur through a variety of modes such as plastic or visco-plastic deformation of material, wear of material through breakdown of lubrication, entry of debris into contact, crack propagation etc.

Durability of interacting surfaces can be enhanced through several methods. According to Ludema [126] there are three methods of modifying surfaces which are surface treatment [127], surface modification [128] and surface coating [129]. The effects of atmospheric agents on exposed metal alloy surfaces are studied and their durability has been evaluated in Refs. [130–132]. According to Bhushan [133], wear reduction can occur through non-uniformly tall, mean pressure of rounded peaks when the contacts should be lesser than the yield strength of the softer material in the contact. Surface hardening can also improve durability [134]. The use of suitable greases lubricants can also extend the life of the interacting surfaces [135, 136]. The durability of grease can be further enhanced through additives [137]. Donnet and Erdemir [138] has presented a review of interfaces with attention to solid lubricants which extend wear life.

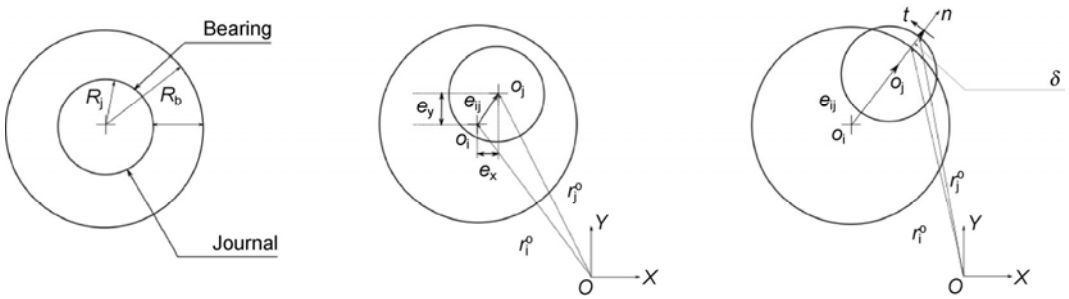


Fig. 15 Representation of the journal bearing with clearance for planar case based on Ref. [98]

The enhancement of wear performance through diamond like carbon (DLC) coatings has also been discussed [138]. The use of diamond like carbon coatings has been presented in Ref. [139] and significantly reduces friction. Results of nano-composite coating, friction and wear analyses in rolling contact have been presented in Ref. [129]. Analysis of interaction at the joint for PEEK and Al 7075 alloys in robotic arms has been presented by Koike et al. [140]. Erdemir [141] has recommended the texturing of surfaces to improve the retention of lubrication and provide superior wear resistance in the contact. The addition of surfactants is also expected to enhance the contact durability by modifying the surface characteristics of the contact [142, 143].

Progress has been made from simple surface hardening to the use of surfactants to enhance wear resistance of contacting surfaces. Continued analysis of the influence of external agents in joint contacts and the influence on dynamics at the contact and investigation of the improvement of wear resistance and friction characteristics of the contact surfaces, surface modification and coatings is required for specialised applications. Furthermore, tribological testing is required to ensure the resilience of the interacting joints. Virtual prototyping and simulation [27, 144], along with the tribological experimentation is necessary for accurate prediction and to enhance the durability of interacting surfaces.

A rapid progress is desired in the modelling techniques and simulations, which bring it asymptotically close to actual physical models. In addition, the increasing reliance on simulation packages and virtual prototyping has been summarised in the forthcoming section.

7 Software packages used in multi-body dynamics

Several commercial software packages are available for the simulation of multibody dynamics [67]. For rigid link multibody mechanics simulation SimMechanics 1st or 2nd generation [53, 145] packages can be used. For joints with clearance packages like ADAMS [25,146, 147], COMPAMM, NEWEUL, DAP3D [148], MUBODYNA [102], RAPID and PARASOLIDS have been used [25] for interference detection. The problem can be formulated by using programming languages such as C/C++, Python or MATLAB m-code, depending on the intended application. However, the process of modelling the mechanics from first principles is often too tedious. Increasing reliance on software simulation packages has been observed with increasing model complexity, shorter project schedules, and higher productivity demand.

8 Conclusions

This paper covers literature in dynamic modelling and friction. It captures the evolution of mechanism and friction dynamics briefly. Key improvements in the area have been identified and presented. The effects of friction models in manipulator dynamics are seldom discussed because of the difficulty of incorporating all the influencing factors into a single model. However, with the improvements in computational and numerical modelling techniques, the frictional dynamics of mechanisms is more effectively analysed.

Rapid evolution of mechanical modelling methods over the latter half of the 20th century includes advances in modelling techniques and computational methods [149]. These enable the modelling and the simulation of complex mechanisms with increased accuracy. The progress from simple kinematics to complex dynamics has been effected through the implementation of several advanced modelling techniques. Improvements in computational capacity have also enhanced the tools available resulting in the reduction of the overall process time.

Friction models have progressed from simple Coulomb, viscous and Stribeck friction models to the more comprehensive dynamic models such as the Leuven model. Some models incorporate wear in the contact by incorporating finite element computational techniques. High degree of non-linearity of the friction model and factors influencing the contact including the surface conditions, material properties, contact conditions and lubricating conditions are among several other factors and the interaction of these factors ensure that the convergence of a single friction modelling equation does not occur. To add to the

complexity, dynamic modelling of mechanisms involves additional uncertainties such as the end effector trajectories [150] and environmental interaction [100]. Therefore, it will be necessary to examine each case individually to determine the important influencing factors in each case.

Progress in dynamic modelling methods with improved friction modelling is envisaged, e.g., a hybrid model that is capable of switching between different regimes. However, the unobservable transition conditions in contact makes this task tedious. Until such a unified model can be derived, smaller unified models addressing specific conditions of contact and frictional force generated in such contacts will be useful. This philosophy is consistent with literature. The influence of environments on contacts determines the life of the mechanism [129]. Improvement of the dynamic modelling techniques and tribo-analysis of the material at the contact is imperative within this context. Further analysis to determine the relationship among dynamics of mechanism, material properties, coefficient of friction and the influence of environments is needed (Fig. 16).

Tribo-testing carried out under differt loads, lubrication conditions and environmental influences would elucidate the dynamics and tribological performance of such contacts. The interaction between any two or more influencing factors may lead to accelerated failure at the joint or an inordinate rise in frictional resistance which are detrimental to the manipulator operation especially when precision positioning while handling heavy loads is required. The data generated thereof can be used to construct a specific but comprehensive model for the above-mentioned factors.

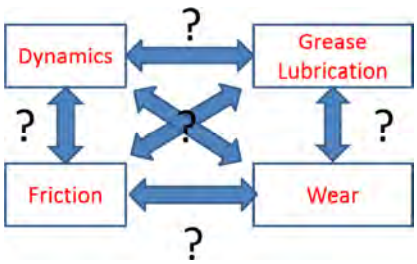


Fig. 16 The relationship between various components of a mechanism requires further study. After several decades of research there is no governing principle owing to the sheer complexity of the phenomenon involved.

References

[1] Saha A, Wahi P, Bhattacharya B. Characterization of friction force and nature of bifurcation from experiments on a single-degree-of-freedom system with friction-induced vibrations. *Tribol Int* **98**:. 220–228 (2016)

[2] Drucker D C. Coulomb friction, plasticity and limit loads. *J Appl Mech* **21**(1):71–74 (1953)

[3] Mukras S M. Analysis and design of planar multibody systems with revolute joint wear. *Dissertations & Theses-Gradworks* **268**(5–6): 643–652 (2009)

[4] Tafazoli S, Lawrence P D, Salcudean S E. Identification of inertial and friction parameters for excavator arms. *IEEE Trans Robot Autom* **15**(5): 966–971 (1999)

[5] Jonker J B, Waiboer R R, Aarts R G K M. Modelling of joint friction in robotic manipulators with gear transmissions. In *Computational Methods in Applied Sciences*. Springer, 2005: 221–243.

[6] Simoni L, Beschi M, Legnani G, Visioli A. Friction modeling with temperature effects for industrial robot manipulators. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015: 3524–3529.

[7] Piedboeuf J-C, de Carufel J, Hurteau R. Friction and stick-slip in robots: Simulation and experimentation. *Multibody Sys Dyn* **4**(4): 341–354 (2000)

[8] Yao B, Tomizuka M. Adaptive control of robot manipulators in constrained motion. In American Control Conference. 1993: 1128–1132.

[9] Nguyen K D, Dankowicz H. Adaptive control of underactuated robots with unmodeled dynamics. *Rob Auton Syst* **64**: 84–99 (2014)

[10] Al-Ashoor R A, Khorasani K, Patel R V, Al-Khalili A J. Adaptive control of flexible joint manipulators. In *1990 IEEE International Conference on Systems, Man, and Cybernetics Conference Proceedings*, 1990: 627–632.

[11] Johnson K L. Contact mechanics. *J Tribol* **108**(4): 464 (1986)

[12] Film thicknesses in lubricated Hertzian contacts (EHL). Part 1: Two-dimensional contacts (line contacts). Engineering Sciences Data Unit (ESDU) 85027, 2010.

[13] Hess D P, Soom A. Friction at a lubricated line contact operating at oscillating sliding velocities. *J Tribol* **112**(1): 147 (1990)

[14] Tyfour W R, Beynon J H, Kapoor A. Deterioration of rolling contact fatigue life of pearlitic rail steel due to dry-wet rolling-sliding line contact. *Wear* 197(1): 255–265 (1996)

[15] Bilandi S T. Identification of frictional effects and structural dynamics for improved control of hydraulic manipulators. *Journal of Crystal Growth* **180**(180): 157–166 (1961)

[16] Flores P, Ambrósio J, Claro J C P, Lankarani H M, Koshy C S. A study on dynamics of mechanical systems including joints with clearance and lubrication. *Mech Mach Theory* **41**(3): 247–261 (2006)

[17] Dupac M, Noroozi S. Dynamic modeling and simulation of a rotating single link flexible robotic manipulator subject to quick stops. *J Mech Eng* **60**(7–8): 475–482 (2014)

[18] Flores P, Ambrósio J, Claro J C P. Dynamic analysis for planar multibody mechanical systems with lubricated joints. *Multibody Syst Dyn* 12(1): 47–74 (2004)

[19] Askari E, Flores P, Dabirrahmani D, Appleyard R. Dynamic modeling and analysis of wear in spatial hard-on-hard couple hip replacements using multibody systems methodologies. *Nonlinear Dyn* **82**(1–2): 1039–1058 (2015)

[20] Flores P. Modeling and simulation of wear in revolute clearance joints in multibody systems. *Mech Mach Theory* **44**(6): 1211–1222 (2009)

[21] Flores P. A parametric study on the dynamic response of planar multibody systems with multiple clearance joints. *Nonlinear Dyn* **61**(4): 633–653 (2010)

[22] Flores P. Compliant contact force approach for forward dynamic modeling and analysis of biomechanical systems. *Procedia IUTAM* **2**: 58–67 (2011)

[23] Marques F, Flores P, Lankarani H M. On the frictional contacts in multibody system dynamics. In *ECCOMAS Themat. Conf. Multibody Dyn.*, 2015: 565–576.

[24] Lankarani H M, Nikravesh P E. Canonical impulse-momentum equations for impact analysis of multibody systems. *J Mech Des* **114**(1): 180–186 (1992)

[25] Koshy C S, Flores P, Lankarani H M. Study of the effect of contact force model on the dynamic response of mechanical systems with dry clearance joints: computational and experimental approaches. *Nonlinear Dyn* **73**(1–2): 325–338 (2013)

[26] Marques F, Flores P, Lankarani H M. Study of friction force model parameters in multibody dynamics. In *The 4th Joint International Conference on Multibody System Dynamics*, 2016.

[27] Mukras S, Kim N H, Mauntler N A, Schmitz T L, Sawyer W G. Analysis of planar multibody systems with revolute joint wear. *Wear* **268**(5–6): 643–652 (2010)

[28] Mukras S, Kim N H, Sawyer W G, Jackson D B, Bergquist L W. Numerical integration schemes and parallel computation for wear prediction using finite element method. *Wear* 266(7–8): 822–831 (2009)

[29] Lingard H, Cooke T, Gharaie E. A case study analysis of fatal incidents involving excavators in the Australian construction industry. *Eng Constr Archit Manag* 20(5): 488–504 (2013)

[30] Koivo A J. Kinematics of excavators (backhoes) for transferring surface material. *J Aerosp Eng* **7**(1): 17–32 (1994)

[31] Vähä P K, Skibniewski M J. Dynamic model of excavator. *J Aerosp Eng* **6**(2): 148–158 (1993)

[32] Koivo A J, Thoma M, Kocaoglan E, Andrade-Cetto J. Modeling and control of excavator dynamics during digging operation. *J Aerosp Eng* **9**(1): 10–18 (1996)

[33] Stentz A, Bares J, Singh S, Rowe P. Robotic excavator for autonomous truck loading. *Auton Robots* **7**: 175–186 (1999)

[34] Seward D, Margrave F. LUCIE the robot excavator-design for system safety. In *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference*, 1996: 963–968.

[35] Tafazoli S, Lawrence P D P, Chan D, Bachmann S, de Silva C. Parameter estimation and actuator friction analysis for a mini excavator. In *Proceedings of the International Conference on Robotics and Automation*, 1996: 329–334.

[36] Carlson J, Murphy R R. How UGVs physically fail in the field. *IEEE Trans Robot* **21**(3): 423–437 (2005)

[37] Liu Y, Nejat G. Robotic urban search and rescue: a survey from the control perspective. *J Intell Robot Syst* **72**(2): 147–165 (2013)

[38] Lioy P J, Weisel C P, Millette J R, Eisenreich S, Vallero D, Offenberg J, Buckley B, Turpin B, Zhong M, Cohen M D, Prophete C, Yang I, Stiles R, Chee G, Johnson W, Porcja R, Alimokhtari S, Hale R C, Weschler C, Chen L C. Characterization of the dust/smoke aerosol that settled east of the World Trade Center (WTC) in lower Manhattan after the collapse of the WTC 11 September 2001. *Environ Health Perspect* **110**(7): 703–714 (2002)

[39] Blau P J. Introduction. In *Friction Science and Technology*. CRC Press, 2008: 1–16.

[40] Khan Z A, Saeed A, Gregory O, Ghafoor A. Biodiesel performance within internal combustion engine fuel system─A review. *Tribol Ind* **38**(2): 197–213 (2016)

[41] Khan Z A, Hadfield M. Manufacturing induced residual stress influence on the rolling contact fatigue life performance of lubricated silicon nitride bearing materials. *Mater Des* **28**(10): 2688–2693 (2007)

[42] Khan Z A, Hadfield M, Tobe S, Wang Y. Ceramic rolling elements with ring crack defects─A residual stress approach. *Mater Sci Eng A* **404**(1–2): 221–226 (2005)

[43] Khan Z A, Hadfield M, Tobe S, Wang Y. Residual stress variations during rolling contact fatigue of refrigerant lubricated silicon nitride bearing elements. *Ceram Int* **32**(7): 751–754 (2006)

[44] Khan Z A, Hadfield M, Wang Y. Pressurised chamber design for conducting rolling contact experiments with liquid refrigerant lubrication. *Mater Des* **26**(8): 680–689 (2005)

[45] Armstrong-Hélouvry B, Dupont P, De Wit C C. A survey of models, analysis tools and compensations methods for the control of machines with friction. *Automatica* **30**(7): 1083–1138 (1994)

[46] Rahnejat H. Multi-body dynamics: Historical evolution and application. *Proc Inst Mech Eng Part C J Mech Eng Sci* **214**(1): 149–173 (2000)

[47] Schiehlen W. Multibody system dynamics: Roots and perspectives. *Multibody Syst Dyn* **1**(2): 149–188 (1997)

[48] Siciliano B, Khatib O. *Springer Handbook of Robotics, 1st ed*. Springer-Verlag Berlin Heidelberg, 2008.

[49] Uicker J J, Pennock G R, Shigley J E, McCarthy J M. Theory of machines and mechanisms. *J Mech Design* **125**(3): 650 (2003)

[50] Groover Jr M P, Zimmers Jr E W. *CAD/Cam: Computer-Aided Design and Manufacturing, 1st ed*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1997.

[51] Farhat N, Mata V, Diaz-Rodriguez M. Dynamic simulation of a parallel robot: Coulomb friction and stick slip in robot joints. *Robotica* **28**: 35–45 (2014)

[52] Hachkowski M R, Peterson L D, Lake M S. Friction model of a revolute joint for a precision deployable spacecraft structure. *J Spacecr Rockets* **36**(4): 591–598 (1999)

[53] Chacko V, Yu H. Multi-body simulation methods for rigid manipulators. *Eurathlon/SHERPA*, 2015.

[54] Donald B, Spong M W. Chapter 3: Forward Kinematics: The Denavit-Hartenberg Convention.

[55] Vähä P K, Skibniewski M J, Koivo A J. Excavator dynamics and effect of soil on digging. In *Proc. 8th ISARC, Stuttgart, Ger.*, 1991: 297–306.

[56] Baiges-Valentin and Ivan J. Baiges-Valentin. Dynamic modeling of parallel manipulators. University of Florida, 1996.

[57] Visioli A, Legnani G. On the trajectory tracking control of industrial SCARA robot manipulators. *IEEE Trans Ind Electron* **49**(1): 224–232 (2002)

[58] Patel B P, Prajapati J M. Dynamics of mini hydraulic backhoe excavator: A Lagrange-Euler (LE) approach. *Int J Mech Aerospace Ind Mechatron Manuf Eng* 8(1): 202–211 (2014)

[59] Korayem M H, Shafei A M. Application of recursive Gibbs–Appell formulation in deriving the equations of motion of N-viscoelastic robotic manipulators in 3D space using Timoshenko Beam Theory. *Acta Astronaut* **83**: 273–294 (2013)

[60] Piedboeuf J C. Kane's equations or Jourdain's principle? In *Proc. 36th Midwest Symp Circuits Syst*, 1993: 1471–1474.

[61] Šalinic S. Determination of joint reaction forces in a symbolic form in rigid multibody systems. *Mech Mach Theory* **46**(11): 1796–1810 (2011)

[62] Obermayr M, Vrettos C, Eberhard P, Däuwel T. A discrete element model and its experimental validation for the prediction of draft forces in cohesive soil. *J Terramechanics* 53: 93–104 (2014)

[63] Yousefi Moghaddam R, Kotchon A, Lipsett M G. Method and apparatus for on-line estimation of soil parameters during excavation. *J Terramechanics* **49**(3–4): 173–181 (2012)

[64] Perumpral J V, Grisso R D, Desai C S, Perumpral J V, Grisso R D, Desai C S. A soil-tool model based on limit equilibrium analysis. *Trans ASAE* 26(4): 0991–0995 (1983)

[65] Shmulevich I. State of the art modeling of soil-tillage interaction using discrete element method. *Soil Tillage Res* **111**(1): 41–53 (2010)

[66] Hollerbach J M. A recursive lagrangian formulation of manipulator dynamics. 1980.

[67] Flores P, Ambrósio J, Claro J C P, Lankarani H M. Multibody systems formulation. *Lecture Notes in Applied & Computational Mechanics* **34**: 47–66 (2008)

[68] Olsson H, Åström K J, Canudas de Wit C, Gäfvert M, Lischinsky P. Friction models and friction compensation. *European Journal of Control* **4**(3): 176–195 (1998)

[69] Sohail T, Malik A M. Measuring Coulomb and viscous friction in revolute joint. *NUST J Eng Sci* **1**(1): 112–116 (2008)

[70] Berger E. Friction modeling for dynamic system simulation. *Appl Mech Rev* 55(6): 535 (2002)

[71] Harrison J A, Gao G, Schall J D, Knippenberg M T, Mikulski P T. Friction between solids. *Philos Trans A Math Phys Eng Sci* **366**: 1469–1495 (2008)

[72] Swevers J, Al-Bender F, Ganseman C G, Prajogo T. An integrated friction model structure with improved presliding behavior for accurate friction compensation. *IEEE Trans Automat Contr* **45**(4): 675–686 (2000)

[73] Le Tien L, Albu-Schaeffer A, De Luca A, Hirzinger G. Friction observer and compensation for control of robots with joint torque measurement. In *2008 IEEE/RSJ Int. Conf. Intell. Robot. Syst. IROS*, 2008: 3789–3795.

[74] Archard J F. Contact and rubbing of flat surfaces. *J Appl Phys* **24**(8): 981–988 (1953)

[75] Jin Z, Fisher J. Tribology in joint replacement. *Joint Replace Technol* **2008**: 31–55 (2008)

[76] Tabor D. Friction─The present state of our understanding. *J Lubr Technol* **103**(2): 169–179 (1981)

[77] van Geffen V. A study of friction models and friction compensation. 2009.

[78] Lischinsky P, Canudas de Wit C, Morel G. Friction compensation for an industrial hydraulic robot. *IEEE Control Syst Mag* **19**(1): 25–32 (1999)

[79] Dahl P R. A solid friction model. Aerospace Corp El Segundo CA, No. TOR-0158 (3107-18)-1, 1968.

[80] Piatkowski T. Dahl and LuGre dynamic friction models─The analysis of selected properties. *Mech Mach Theory* **73**: 91–100 (2014)

[81] Karnopp D. Computer simulation of stick-slip friction in mechanical dynamic systems. *J Dyn Syst Meas Control* **107**(1): 100–103 (1985)

[82] Kikuuwe R, Takesue N, Sano A, Mochiyama H, Fujimoto H. Fixed-step friction simulation: From classical coulomb model to modern continuous models. In *2005 IEEE/RSJ Int. Conf. Intell. Robot. Syst. IROS*, 2005: 3910–3917.

[83] Kikuuwe R, Takesue N, Sano A, Mochiyama H, Fujimoto H. Admittance and impedance representations of friction based on implicit Euler integration. *IEEE Trans Robot* **22**(6): 1176–1188 (2006)

[84] Awrejcewicz J, Kudra G. Regular and chaotic dynamics of three coupled physical pendulums. In *the International Conference of Applied Mechanics*, Sacam, 2000.

[85] Wojewoda J, Stefański A, Wiercigroch M, Kapitaniak T. Hysteretic effects of dry friction: modelling and experimental studies. *Philos Trans R Soc London A Math Phys Eng Sci* **366**(1866): 747–765 (2008)

[86] Romano R A, Garcia C. Karnopp friction model identification for a real control valve. *IFAC Proc* **17**(1): 14906–14911 (2008)

[87] Bliman P A, Sorine M. Easy-to-use realistic dry friction models for automatic control. In *Proceedings of 3rd European Control Conference*, 1995: 267–272.

[88] Gafvert M. Comparisons of two dynamic friction models. In *IEEE International Conference on Control Applications*, 1997: 386–391.

[89] Freidovich L, Robertsson A, Shiriaev A, Johansson R. LuGre-model-based friction compensation. *IEEE Trans Control Syst Technol* **18**(1): 194–200 (2010)

[90] Lampaert V, Swevers J, Al-Bender F. Modification of the Leuven integrated friction model structure. *IEEE Trans Automat Contr* **47**(4): 683–687 (2002)

[91] The Mathworks Inc. UK. Loaded-contact rotational friction. *Mathworks Doc* 2016: 2–4.

[92] The Mathworks Inc. UK. Translational friction. *Mathworks Documentation*, 2015.

[93] Bona B, Indri M. Friction compensation in robotics: An overview. In *Proc. 44th IEEE Conf Decis Control Eur Control Conf CDC-ECC '05*, 2005: 4360–4367.

[94] De Wit C C, Olsson H, Astrom K J, Lischinsky P. A new model for control of systems with friction. *IEEE Trans Automat Contr* **40**(3): 419–425 (1995)

[95] Haessig D A, Friedland B. On the modeling and simulation of friction. In *Am Control Conf*, 1990: 1256–1261.

[96] Dubowsky S, Freudenstein F. Dynamic analysis of mechanical systems with clearances—Part 1: Formation of dynamic model. *J Eng Ind* **93**(1) 305 (1971)

[97] Bauchau O A, Rodriguez J. Modeling of joints with clearance in flexible multibody systems. *Int J Solids Struct* **39**(1): 41–63 (2001)

[98] Bai Z F, Zhao Y. A hybrid contact force model of revolute joint with clearance for planar mechanical systems. *Int J Non Linear Mech* **48**: 15–36 (2013)

[99] Flores P, Ambrósio J, Claro J C P, Lankarani H M, Koshy C S. Lubricated revolute joints in rigid multibody systems. *Nonlinear Dyn* **56**(3): 277–295 (2009)

[100] Wang Y. Simulation of mechanical systems with multiple frictional contacts. *J Mech Des* **116**(2): 571 (1994)

[101] Haines R S. Survey: 2-dimensional motion and impact at revolute joints. *Mech Mach Theory* **15**(5): 361–370 (1980)

[102] Machado M, Costa J, Seabra E, Flores P. The effect of the lubricated revolute joint parameters and hydrodynamic force models on the dynamic response of planar multibody systems. *Nonlinear Dyn* **69**(1–2): 635–654 (2012)

[103] Dubowsky S, Freudenstein F. Dynamic analysis of mechanical systems with clearances- Part2: Dynamic response. *J Eng Ind* **93**(1): 305–309 (1971)

[104] Hunt K H, Crossley F R E. Coefficient of restitution interpreted as damping in vibroimpact. *J Appl Mech* **42**(2): 440 (1975)

[105] Lankarani H M, Nikravesh P E. A contact force model with hysteresis damping for impact analysis of multibody systems. *J Mech Des* **112**(3): 369–376 (1990)

[106] Rhee J, Akay A. Dynamic response of revolute joint with clearance. *Mech Mach Theory* **31**(1): 121–134 (1996)

[107] Pereira M S, Nikravesh P E. Impact dynamics of multibody systems with frictional contact using joint coordinates and canonical equations of motion. *Nonlinear Dyn* 9(1–2): 53–71 (1996)

[108] Machado M, Moreira P, Flores P, Lankarani H M. Compliant contact force models in multibody dynamics: Evolution of the Hertz contact theory. *Mech Mach Theory* **53**: 99–121 (2012)

[109] Flores P, Koshy C S, Lankarani H M, Ambrósio J, Claro J C P. Numerical and experimental investigation on multibody systems with revolute clearance joints. *Nonlinear Dyn* 65(4): 383–398 (2011)

[110] Ravn P. A continuous analysis method for planar multibody systems with joint clearance. *Multibody Syst Dyn* **2**(1): 1–24 (1998)

[111] Dupac M. Dynamical analysis of a constrained flexible extensible link with rigid support and clearance. *J Theoret Appl Mech* **52**(3): 665-676 (2014)

[112] Azimi Olyaei A, Ghazavi M R. Stabilizing slider-crank mechanism with clearance joints. *Mech Mach Theory* **53**: 17–29 (2012)

[113] Zhao B, Zhang Z-N, Fang C, Dai X-D, Xie Y-B. Modeling and analysis of planar multibody system with mixed lubricated revolute joint. *Tribol Int* **98**: 229–241 (2016)

[114] Furuhashi T, Morita N, Matsuura M. Research on dynamics of four-bar linkage with clearances at turning pairs: 1st report, general theory using continuous contact model. *Bull JSME* **21**(153): 518–523 (1978)

[115] Farahanchi F, Shaw S W. Chaotic and periodic dynamics of a slider-crank mechanism with slider clearance. *J Sound Vib* **177**(3): 307–324 (1994)

[116] Khulief Y A, Shabana A A. A continuous force model for the impact analysis of flexible multibody systems. *Mech Mach Theory* **22**(3): 213–224 (1987)

[117] Dubowsky S, Gardner T N. Dynamic interactions of link elasticity and clearance connections in planar mechanical systems. *J Eng Ind* **97**(2): 652–661 (1975)

[118] Earles S W E, Wu C L S. Motion analysis of a rigid link mechanism with clearance at a bearing using Lagrangian mechanics and digital computation. *J Mech* 83–89 (1973)

[119] Yaqubi S, Dardel M, Daniali H M, Hassan M, Ghasemi M H. Modeling and control of crank-slider mechanism with multiple clearance joints. *Multibody Syst Dyn* **36**(2): 143–167 (2016)

[120] Li P, Chen W, Li D, Yu R. A novel transition model for lubricated revolute joints in planar multibody systems. *Multibody Syst Dyn* **36**(3): 279–294 (2016)

[121] Flores P, Leine R, Glocker C. Modeling and analysis of planar rigid multibody systems with translational clearance joints based on the non-smooth dynamics approach. *Multibody Syst Dyn* **23**(2): 165–190 (2010)

[122] Flores P, Ambrósio J. Revolute joints with clearance in multibody systems. *Comput Struct* **82**(17–19): 1359–1369 (2004)

[123] Chytka T M, Brown R W, Shih A T, Reeves J D, Dempsey J A. An integrated approach to life cycle analysis. In *11th AIAA/ISSMO Multidiscip Anal Optim Conf*, 2006: 1360–1374.

[124] Carlson J, Murphy R R. Reliability analysis of mobile robots. In *2003 IEEE Int Conf Robot Autom*, 2003: 274–281.

[125] Wang Z. Design study of an earthquake rescue robot. Durham University, 2004.

[126] Ludema K C. *Friction, Wear, Lubrication - A Textbook in Tribology*. CRC Press, 1996.

[127] Chang S-H, Huang K-T, Wang Y-H. Effects of thermal erosion and wear resistance on AISI H13 tool steel by various surface treatments. *Mater Trans* **53**(4): 745–751 (2012)

[128] Wakuda M, Yamauchi Y, Kanzaki S, Yasuda Y. Effect of surface texturing on friction reduction between ceramic and steel materials under lubricated sliding contact. *Wear* **254**(3–4): 356–363 (2003)

[129] Bajwa R S, Khan Z A, Bakolas V, Braun W. Water-lubricated Ni-based composite (Ni–Al_2O_3, Ni–SiC and Ni–ZrO_2) thin film coatings for industrial applications. *Acta Metall Sin* **29**(1): 8–16 (2016)

[130] Saeed A. Sustainable methodology of conserving historic military vehicles. Bournemouth University, 2013.

[131] Nazir M H, Khan Z A, Stokes K R. A holistic mathematical modelling and simulation for cathodic delamination mechanism—A novel and an efficient approach. *J Adhes Sci Technol* **29**(22): 2475–2513 (2015)

[132] Nazir M H, Khan Z A, Stokes K R. Optimisation of interface roughness and coating thickness to maximise coating–substrate adhesion – a failure prediction and reliability assessment modelling. *J Adhes Sci Technol* **29**(14): 1415–1445 (2015)

[133] Bhushan B. Contact mechanics of rough surfaces in tribology: multiple asperity contact. *Tribol Lett* **4**(1): 1–35 (1998)

[134] Totik Y, Sadeler R, Altun H, Gavgali M. The effects of induction hardening on wear properties of AISI 4140 steel in dry sliding conditions. *Mater Des* **24**(1): 25–30 (2003)

[135] Types and applications of lubricating greases. In *Tribology Series*, 1992: 664–696.

[136] De Laurentis N, Kadiric A, Lugt P, Cann P M E. The influence of bearing grease composition on friction in rolling/sliding concentrated contacts. *Tribol Int* 94: 624–632 (2016)

[137] Cann P M E. Understanding grease lubrication. *Tribol Ser* **31**: 573–581 (1996)

[138] Donnet C, Erdemir A. Historical developments and new trends in tribological and solid lubricant coatings. *Surf Coat Technol* **180–181**: 76–84 (2004)

[139] Erdemir A, Eryilmaz O L, Fenske G. Synthesis of diamondlike carbon films with superlow friction and wear properties. *J Vac Sci Technol A* **18**(4): 1987-1992 (2000)

[140] Koike H, Kanemasu K, Itakura K,. Wear and transmission error between PEEK bush and 7075 aluminium alloy cam plate components in robot joints. *Appl Mech Mater* **307**: 3–8 (2013)

[141] Erdemir A. Review of engineered tribological interfaces for improved boundary lubrication. *Tribol Int* **38**(3): 249–256 (2005)

[142] Sahoo P, Das S K. Tribology of electroless nickel coatings—A review. *Mater Des* 32(4): 1760–1775 (2011)

[143] Zhang J, Meng Y, Tian Y, Zhang X. Effect of concentration and addition of ions on the adsorption of sodium dodecyl sulfate on stainless steel surface in aqueous solutions. *Colloids Surfaces A Physicochem Eng Asp* **484**: 408–415 (2015)

[144] Choi S H, Cheung H H. A versatile virtual prototyping system for rapid product development. *Comput Ind* **59**(5): 477–488 (2008)

[145] Schlotter M. Multibody system simulation with simmechanics. University of Canterbury, 2003.

[146] Tuominen M. Multibody simulations as a product development tool: Introduction to ADAMS and two analyses. Aalto University School of Engineering, 2015.

[147] Ben Abdallah M A, Khemili I, Aifaoui N. Numerical investigation of a flexible slider-crank mechanism with multijoints with clearance. *Multibody System Dynamics* **38**(2): 173-199 (2016)

[148] Ambrósio J, Verissimo P. Improved bushing models for general multibody systems and vehicle dynamics. *Multibody Syst Dyn* **22**(4): 341–365 (2009)

[149] Khurshid A, Zulfiqar K, Chacko V, Ghafoor A, Malik M A, Ayaz Y. Modelling and simulation of a manipulator with stable viscoelastic grasping incorporating friction. *Tribol Ind* **38**(4): 559–574 (2016)

[150] Chacko V, Yu H, Cang S, Vladareanu L. State of the art in excavators. In *International Conference on Advanced Mechatronic Systems, ICAMechS*, 2014: 481–488.

Zulfiqar A. KHAN. He is a professor of design, engineering & computing at Bournemouth University, UK. He completed his BEng (Hons) and MEng (Hons) degrees in mechanical engineering and a PhD degree in tribology. He has previously worked in automotive, aircraft industries and GTZ United Nations energy savings technologies for several years. His research expertise includes corrosion fatigue, rolling contact fatigue, tribology and nanocoating with over 120 publications and a handbook on pressure vessel steel. He is currently leading a research portfolio of six industrial projects within NanoCorr, Energy & Modelling (NCEM) at Bournemouth University.

This review is part of a wider multidisciplinary research within NCEM led by Professor Khan and focuses on dynamic modelling and simulation to optimise friction through surface engineering within interacting joints applied in harsh environments. Authors are delighted to have put this work together to be available to scientists, academics, researchers, professionals and students to engage in the design for durability and reliability through a multidisciplinary approach to bring about societal, economic and environmental benefits.

Friction **5**(1): 23–31 (2017)
DOI 10.1007/s40544-016-0128-4
ISSN 2223-7690
CN 10-1237/TH
RESEARCH ARTICLE

Computational investigation of the lubrication behaviors of dioxides and disulfides of molybdenum and tungsten in vacuum

Jingyan NIAN[1], Liwei CHEN[2], Zhiguang GUO[1,2,*], Weimin LIU[1]

[1] *State Key Laboratory of Solid Lubrication, Lanzhou Institute of Chemical Physics, Chinese Academy of Sciences, Lanzhou 730000, China*

[2] *Hubei Collaborative Innovation Centre for Advanced Organic Chemical Materials and Ministry of Education Key Laboratory for the Green Preparation and Application of Functional Materials, Hubei University, Wuhan 430062, China*

Received: 18 September 2016 / Revised: 02 November 2016 / Accepted: 03 November 2016

Abstract: Lamellar compounds such as the disulfides of molybdenum and tungsten are widely used as additives in lubricant oils or as solid lubricants in aerospace industries. The dioxides of these two transition metals have identical microstructures with those of the disulfides. The differences in the lubrication behaviors of disulfide and dioxides were investigated theoretically. Tungsten dioxide and molybdenum dioxide exhibit higher bond strengths at the interface and lower interlayer interactions than those of the disulfides which indicates their superlubricity. Furthermore, the topography of the electron density of the single layer nanostructure determined their sliding potential barrier; the dioxides showed a weaker electronic cloud distribution between the two neighboring oxygen atoms, which facilitated the oxygen atoms of the counterpart to go through. For commensurate friction, the dioxides exhibited nearly the same value of friction work, and same was the case for the disulfides. The lower positive value of friction work for the dioxides confirmed their improved lubricity than the disulfides and the higher mechanical strength of the bulk dioxides demonstrated that they are excellent solid lubricants in vacuum.

Keywords: solid lubricant; superlubricity; first-principles; molecular dynamics; disulfides; dioxides

1 Introduction

Molybdenum disulfide (MoS_2), a well-known lamellar compound has been widely used in the aerospace industry and as an additive in lubricant due to its low friction coefficient [1−3]. Stefanov et al. [4] studied the lubrication mechanism of MoS_2 nanotubes at the atomic scale using molecular dynamics and found that the lubrication process at high loads is not determined by the ball-bearing effect; the nanotube is compressed to form a nanosheet resulting in excellent lubrication performance. Onodera et al. [5] studied the friction process between the interlayers of MoS_2 nanosheet using a hybrid quantum chemical/classical molecular dynamics method. The results indicated that the strong Coulombic force of repulsion between the interlayers of the nanosheets leads to the low friction; if the MoS_2 nanosheets are adsorbed on a substrate with electron-donating property, the friction will be further reduced due to the larger Coulombic force of repulsion. The dynamic friction of two-layer MoS_2 nanosheets at the atomic level was also theoretically investigated by the molecular dynamics method by Onodera et al. [6]. It was demonstrated that incommensurate structures showed extremely low friction due to the cancellation of the atomic forces in the sliding direction. On the other hand, in commensurate structures, the atomic forces were formed by the approaching of atoms in sliding part with the atoms in its counterpart, and all the atomic forces acted in the same direction, leading to the accumulation of lateral forces and higher friction. Wang et al. [7] investigated the friction property of in-plane stretch and compression of MoS_2 sheets and reported that the in-plane compression can also reduce

* Corresponding author: Zhiguang GUO, E-mail: zguo@licp.cas.cn

the friction due to the decrease in the potential energy of sliding. Levita et al. [8] studied the influence of load and interlayer orientation on the friction of 2H MoS_2 using the *ab initio* method and the results indicated that the electrostatic interaction progressively affects the sliding motion at higher loads.

In addition to MoS_2, tungsten disulfide (WS_2) is also an important solid lubrication material that has been used for various tribological applications. However, theoretical studies on the friction behavior of WS_2 have been rarely carried out to quantitatively reveal the differences in the lubrication behaviors of MoS_2 and WS_2. This motivated us to study the lubrication behavior of WS_2. Furthermore, the superlubricity of tungsten dioxide (WO_2) was demonstrated by Cahangirov et al. [9] using the "Prandtl-Tomlinson model" [10, 11]. Therefore, it would be interesting to study the lubrication property of molybdenum dioxide.

In the present work, the differences in the lubrication behaviors of MoS_2, WS_2, MoO_2, and WO_2 solid lubricants were investigated using the first principles theory and molecular dynamic simulations.

2 Computational details

All the solid lubricant materials mentioned above consist of hexagonal unit cells, formed by stacking X-M-X (M = Mo, W; X = S, O) trilayers; hexagonal unit cells are weakly bonded to each other by the van der Waals forces. It is to be noted that the stable existence of these single layer solid materials has been demonstrated through optimization of the geometries and phonon dispersion calculations based on first-principles [12−14].

2.1 Intrinsic properties of single layer nanostructures

Considering single layer nanostructures as the research models, all the density functional theory plane-wave pseudopotential calculations were carried out using CASTEP [15, 16] which is a state-of-the-art quantum mechanics-based program. The intrinsic properties such as electronic structures and density of states were determined by LDA-CA-PZ (Ceperly and Alder's local density approximation [17, 18]) methods. The kinetic energy cutoff for the plane-wave basis set was 300 eV, the convergence for energy was chosen as 5.0×10^{-6} eV between two steps, the maximum force allowed on each atom was less than 10^{-2} eV/Å, the maximum stress and the displacement were 0.02 GPa and 5.0×10^{-7} nm respectively, the Brillouin zone was sampled by (12 × 12 × 2), and the ionic species were described by ultrasoft pseudopotentials.

2.2 Interlayer sliding interaction

The classic first-principle theory of atomic-scale friction proposed by Zhong and Tomanek [19] was used in the present study. Taking MoO_2 (2 × 2 supercells) for instance, the top-view of the calculation scheme is shown in Fig. 1. Ortmann, Bechstedt, and Schmid semiempirical dispersion-correction [20, 21] was added due to the van der Waals force between the interlayers. A vacuum region of 1.5 nm thickness was added in the [001] direction in order to avoid interaction between two single layers.

The interlayer interaction energies of two atomic-scale slabs were calculated by Eq. (1).

$$E_{\text{adsorption}} = E_{\text{Total}} - E_{\text{Slab1}} - E_{\text{Slab2}} \quad (1)$$

$E_{\text{adsorption}}$ and E_{total} represent the interaction energy and total energy of two MoO_2 slabs (2 × 2 supercells) at different relative positions respectively. E_{slab1} and E_{slab2} are the energies of two slabs (2 × 2 supercells).

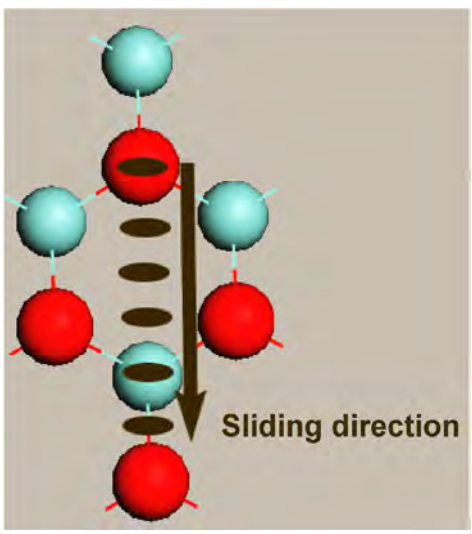


Fig. 1 Schematic diagram (top view) showing the relative sliding between two MoO_2 (2 × 2 × 1) nanostructures. The ellipses indicate the relative positions and the arrow indicates the sliding direction. The red spheres represent the oxygen atoms and the light blue spheres represent the transition metal atoms.

Considering the direction of the arrow shown in Fig. 1 as the sliding direction, x, the position-dependent parts of the potential energy, V can be expressed as shown in Eq. (2).

$$V(x, f_{ext}) = E_{adsorption}(x, z(x, f_{ext})) + f_{ext}z(x, f_{ext}) - V_0(f_{ext}) \quad (2)$$

Here, V_0 is the minimum potential energy corresponding to the O atom located at the hollow site relative to the other slab. In the case of atomic stick-slip motion, the conservative force results in an increase in the potential energy along the sliding path. This process is described by Eq. (3).

$$\Delta V_{max}(f_{ext}) = V_{max}(f_{ext}) - V_{min}(f_{ext}) \quad (3)$$

Therefore, the lateral position-dependent friction force, f_x can be given by Eq. (4),

$$\langle F \rangle = \frac{\Delta V_{max}}{\Delta x} \quad (4)$$

From Eq. (4), it is clear that the lubrication performance is associated with the potential barriers in the sliding path with minimum energy. The lower the value of the potential barrier, the better is the lubrication performance.

$$f_{ext} = -\partial E_{interaction}(z)/\partial z \quad (5)$$

According to Eq. (5), decreasing the interlayer distance is similar to applying a normal load. In the present system, the interlayer distance was decreased to 0.5 Å. It is to be noted that, the contour plots of the potential energy surfaces as functions of the different relative lateral positions of the two layers were constructed using the spline interpolation method [8].

2.3 Mechanical properties of bulk materials

The mechanical properties such as bulk modulus, Young's modulus, shear modulus, compressibility and in-plane Poisson's ratio of the bulk solid lubricant materials were also calculated using the first-principles theory. The shear modulus, G can be calculated using Eq. (6) [22],

$$G = \frac{1}{30}(C_{11} + C_{12} + 2C_{33} - 4C_{13} + 12C_{44} + 12C_{66}) \quad (6)$$

Here, C_{ij} and C_{ii} are the elastic constants, i and j indicate the subscripts.

2.4 Molecular dynamics simulation of commensurate sliding

It is well known that commensurate surfaces usually appear in most true friction course, and thus molecular dynamic simulations were carried out by using Forcite modulus in Material Studio 6.0, Accelrys Software Inc. The commensurate friction models are shown in Fig. 2. The system temperature was set to 1 K (−272.15°C). The universal force field **[**23**]** which is a purely diagonal and harmonic force field was assigned to the transition metals and the light elements. The bond stretching was described by a harmonic term, the angle bending by a three-term Fourier cosine expansion, and the torsion and the inversion were described by Fourier cosine expansion terms. The van der Waals interactions were described by the Lennard-Jones potential and the electrostatic interactions were described by atomic monopoles and a screened Coulombic term. The sliding velocity, v, was 0.05 Å/ps, and a complete stick-slip cycle was taken as the sliding distance. The positive friction work, $W_{friction}$ of the stick stage was used to evaluate the friction performance and calculated using Eq. (7).

$$W_{friction} = \int_0^t F \cdot v \mathrm{d}t \quad (7)$$

The value of F is obtained by summing the lateral forces acting on the atoms at the interface.

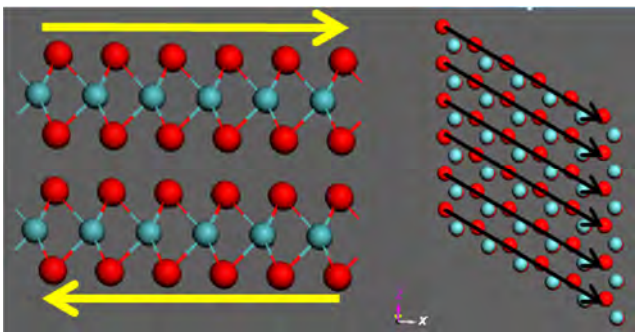


Fig. 2 Schematic showing the relative sliding between two single layer MoO_2 (6 × 6 × 1) nanostructures; arrows show the sliding directions. The red spheres represent the oxygen atoms and the light blue spheres represent the transition metal atoms.

3 Results and discussion

First, the intrinsic properties of the single layer dioxides and disulfides of molybdenum and tungsten are considered, which are summarized in Table 1. The more charged light atoms at the interface correspond to a shorter interlayer distance in the commensurate condition, which indicates that the equilibrium interlayer distance is not determined by the electrostatic repulsion. In order to study the dry friction course of the solid lubricant materials at the atomic scale, it is necessary to investigate the intrinsic stiffness and the interlayer interaction. The intrinsic stiffness of a material can be evaluated from the bond stiffness at the interface. The density of states of the single layer nanostructures are shown in Fig. 3, which shows that the valence bands of all the nanostructures are composed of hybrids of s, p, and d orbitals. In addition, the degree of overlap of the p-d orbitals between the transition metal and the light atoms is stronger than

Table 1 Intrinsic property of the single layer dioxides and disulfides of molybdenum and tungsten. The interlayer distance here is the equilibrium interlayer distance between the bottom atoms of the top layer and the top atoms of the bottom layer in a commensurate condition.

System	Lattice constants (Å)	d_{x-x} (Å)	Bond length (Å)	Bond population	Hirshfeld charge	Interlayer distance (Å)
MoO_2	a=b=2.82	d_{O-O}=2.42	2.04	0.75	O: –0.25; Mo: 0.48	2.84
MoS_2	a=b=3.18	d_{S-S}=3.11	2.41	0.97	S: –0.11; Mo: 0.48	3.46
WO_2	a=b=2.84	d_{O-O}=2.45	2.05	0.83	O: –0.20; W: 0.40	2.87
WS_2	a=b=3.19	d_{S-S}=3.13	2.42	1.08	S: –0.06; W: 0.12	3.49

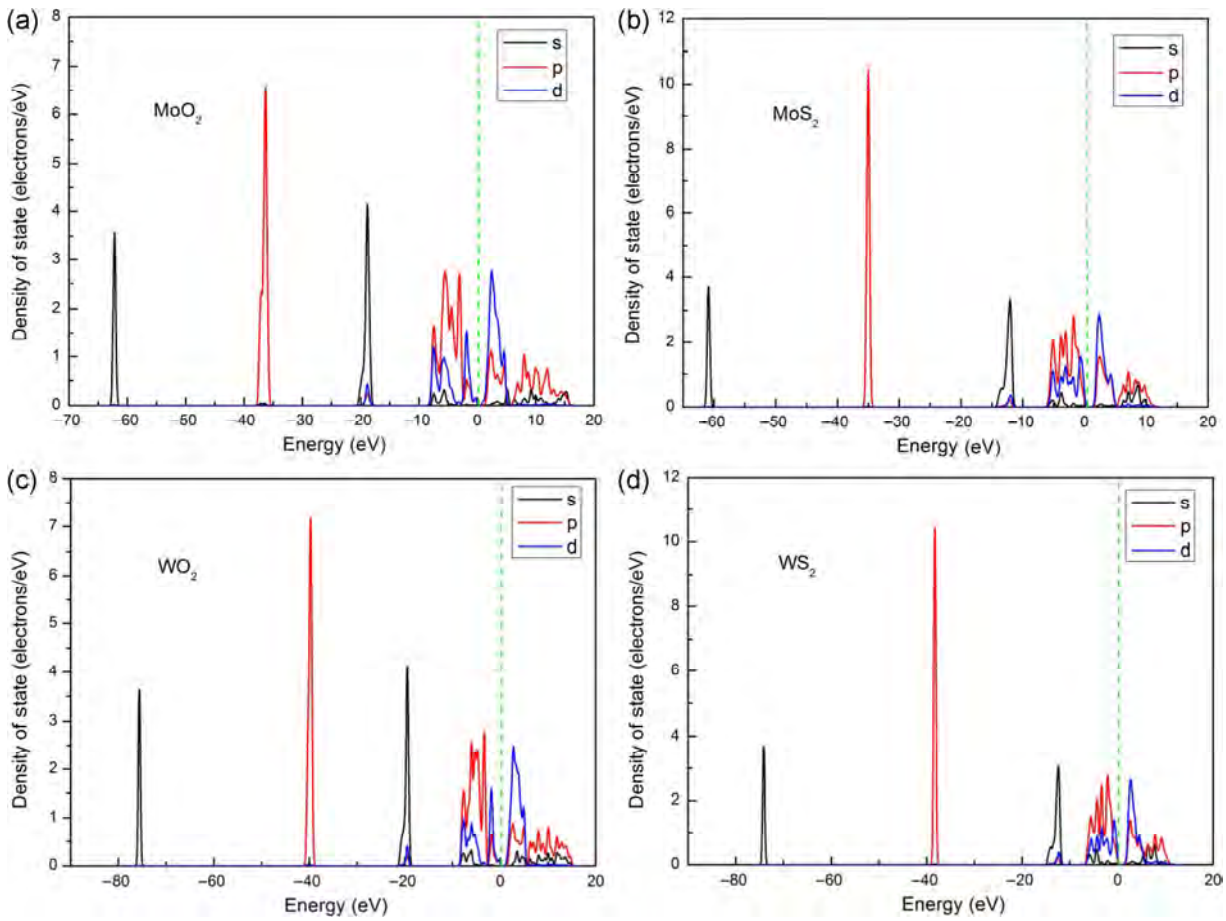


Fig. 3 Density of states distribution of single layer nanostructures of MoO_2, MoS_2, WO_2, and WS_2.

that in the case of the disulfides. Therefore, the M-O (M = Mo, W) bond is shorter than the M-S bond and cannot be broken easily. In other words, the transition metal dioxides exhibit higher bond stiffness at the interface than the disulfides.

Second, the interlayer interaction during the sliding course is taken into consideration. The static potential surfaces of the four solid lubricant materials at their equilibrium interlayer distances are shown in Fig. 4. All the materials exhibit a zigzag minimum energy sliding path. Furthermore, the transition metal dioxides exhibit identical interlayer friction potential barriers, and same is the case with the disulfides. However, the dioxides have a lower interlayer interaction than the disulfides. Therefore, according to the Prandtl-Tomlinson model and the view point of Dag and Ciraci [24], the higher bond stiffness at the interface and the lower interlayer interaction in the dioxides indicate their superlubricity. It is also necessary to investigate the intrinsic electron distribution of the single layer nanostructures. As shown in Fig. 5, the amplitude of the electron density of the transition metal disulfides is higher than that of the dioxides. This explains the lower equilibrium interlayer distance in the case of transition metal dioxides despite their more charged surface. Furthermore, the dioxides have a weaker electronic cloud density between the two neighboring oxygen atoms, which facilitates the oxygen atoms of the counterpart to go through. Therefore, the dioxides exhibit a better lubrication performance than the disulfides due to the sparse electron cloud distribution among the oxygen atoms at the interface. When the interlayer distance decreases to 0.5 Å, as shown in Fig. A1 of the Appendix, the lubrication performance can be arranged in the order WS_2>MoS_2>MoO_2>WO_2 by comparing the values of $\Delta V_{max}(f_{ext})$ and the potential barrier. However, the difference in the frictional potential barrier between the dioxides and the disulfides is only 0.1 eV. Furthermore, the stress in

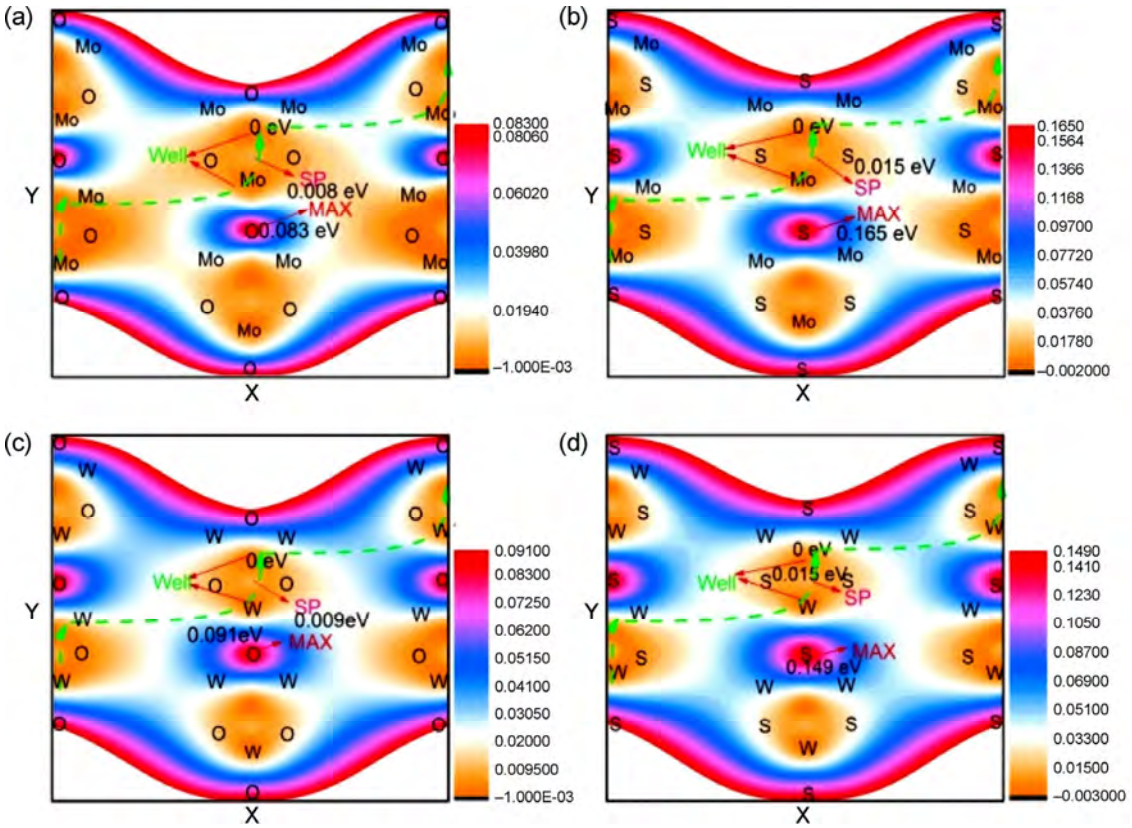


Fig. 4 Static potential surfaces of (a) MoO_2, (b) MoS_2, (c) WO_2 and (d) WS_2 at their equilibrium interlayer distance. Well indicates the minimum interaction energy, SP indicates the saddle point, and MAX indicates the maximum interaction energy. The curved arrows indicate the minimum energy sliding path.

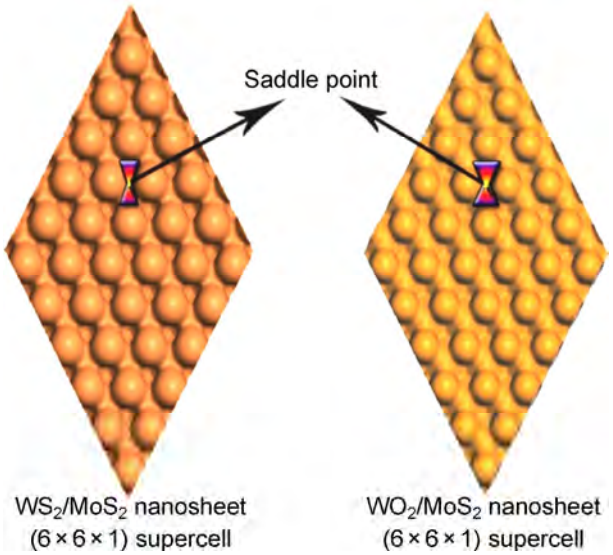


Fig. 5 Topography of the electron distribution of the single layer transition metal disulfide and dioxide nanostructures.

the *z* direction as shown in Fig. A2 of the Appendix clearly indicates that such a small potential energy difference is obtained by almost doubling the normal load. Thus, at the same applied normal load, the dioxides exhibit a better lubrication performance.

Third, the molecular dynamic simulation results are analyzed. As shown in Fig. 6, during the whole stick-slip period, MoS_2 and WS_2 exhibit nearly the same value of friction work, and similar is the case with MoO_2 and WO_2. The lower positive friction work of the dioxides in the stick phase confirms the superlubricity of MoO_2 and WO_2.

Finally, the mechanical properties of the bulk solid lubricant materials are evaluated based on the first principles calculation. The results are summarized in Table 2. The transition metal dioxides exhibit nearly the same values of bulk modulus, Poisson's ratio and compressibility. The mechanical properties of the transition metal disulfides are also similar. However, the dioxides exhibit higher values compared to the disulfides. Therefore, we can conclude that the transitional metal dioxides demonstrate better load bearing capacities than the disulfides. Furthermore, the values of shear modulus and Young's modulus of the disulfides are comparable. It is interesting to note that MoO_2 exhibits higher values of shear modulus and Young's modulus than those of MoS_2 and WS_2; however, these values are lower than those of WO_2. The dioxides therefore exhibit higher mechanical stiffness than the disulfides making them difficult to deform. The higher load bearing capability and mechanical stiffness suggest that MoO_2 and WO_2 are better solid lubricants compared with their disulfides.

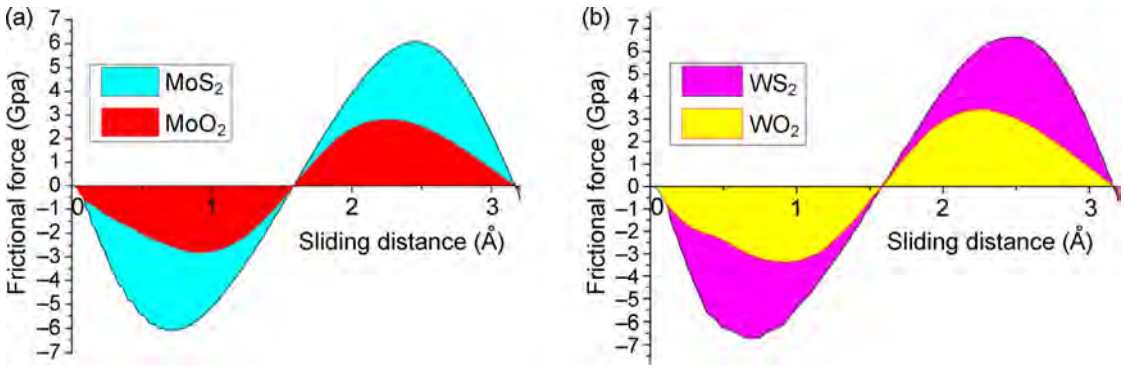


Fig. 6 Interlayer friction work done by (a) MoS_2 and MoO_2, and (b) WS_2 and WO_2, during friction in a commensurate condition.

Table 2 Mechanical properties of bulk MoO_2, MoS_2, WO_2, and WS_2.

System	Bulk modulus (Gpa)	Poisson ratio	Compressibility (1/Gpa)	Shear modulus (Gpa)	Young modulus (Gpa)
MoO_2	66.58	0.35	0.01502	103.6	471.41
MoS_2	58.06	0.23	0.01722	62.89	228.48
WO_2	64.38	0.33	0.01553	131.45	526.08
WS_2	56.26	0.19	0.01778	64.93	249.85

4 Conclusions

The lubrication behaviors of lamellar dioxides and disulfides of molybdenum and tungsten were investigated theoretically and the results are summarized as follows.

(i) The topography of the electron density of the single layer nanostructures determined their sliding potential barrier. The sparser electron cloud distribution between the two neighbor oxygen atoms in the dioxides as compared to that of the disulfides facilitates the oxygen atoms of the counterpart to go through.

(ii) The higher bond strength at the interface, the lower interlayer interaction, and the lower positive friction work during the stick phase indicated the better lubricity of the dioxides than the disulfides.

(iii) The dioxides exhibit nearly similar lubrication performances, so do the disulfides.

(iv) The higher mechanical strength of the bulk dioxide materials indicated that they are better solid lubricants than the disulfides in vacuum.

Acknowledgements

This work is supported by the National Nature Science Foundation of China (Nos. 51522510 and 51675513), the "Top Hundred Talents" Program of Chinese Academy of Sciences and the National Key Basic Research and Development (973) Program of China (2013CB632300) for financial support.

Appendix

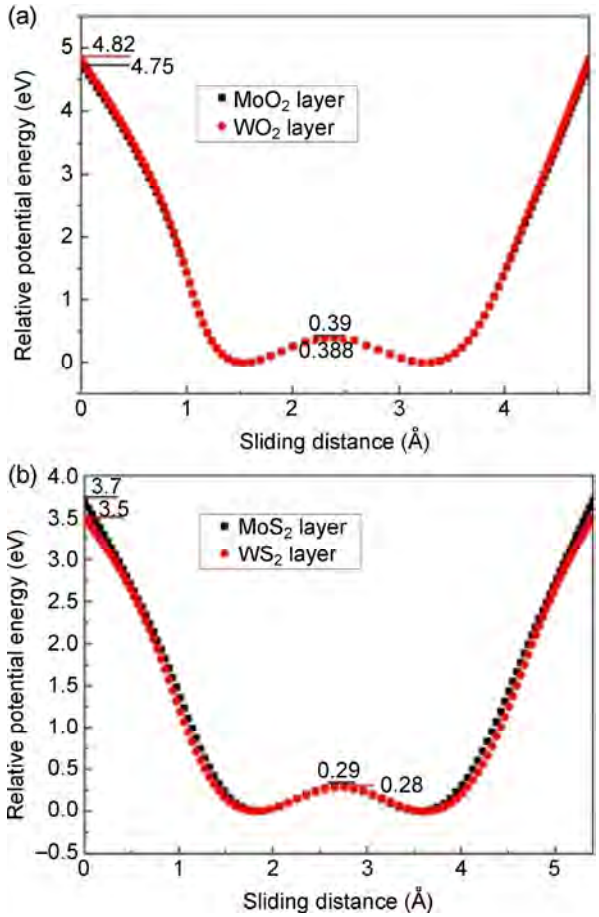


Fig. A1 (a) Relative sliding potential energy curves of transition metal dioxides when interlayer distance decreases 0.5 Å. (b) Relative sliding potential energy curves of transition metal disulfides when interlayer distance decreases 0.5 Å.

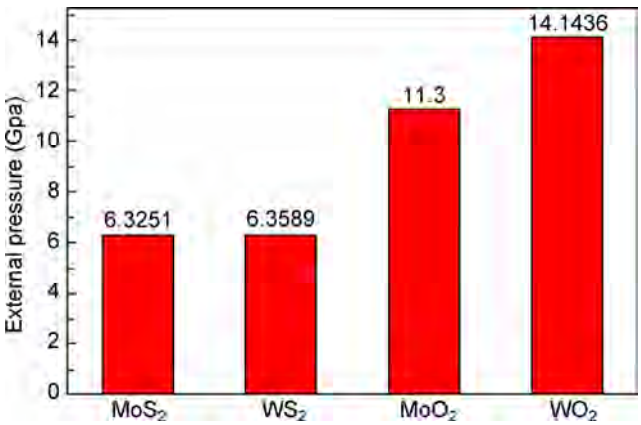


Fig. A2 The normal load is necessary to compress inter-layer distance down by 0.5 Å.

References

[1] Grossiord C, Varlot K J, Martin M, Le Mogne T, Esnouf C, Inoue K. MoS_2 single sheet lubrication by molybdenum dithiocarbamate. *Tribol Int* **31**: 737–743 (1998)

[2] Martin J M, Grossiord C, Le Mogne T, Igarashi J. Transfer films and friction under boundary lubrication. *Wear* **245**: 107–115 (2000)

[3] De Barros Bouchet M I, Martin J M, Le Mogne T, Bilas P, Vacher B, Yamada Y. Mechanisms of MoS_2 formation by MoDTC in presence of ZnDTP: Effect of oxidative degradation. *Wear* **258**: 16430–16450 (2005)

[4] Stefanov M, Enyashin A N, Heine T, Seifert G. Nano-lubrication: How do MoS_2-based nanostructures lubricate? *J Phys Chem C* **112**: 17764–17767 (2008)

[5] Onodera T, Morita Y, Suzuki A, Koyama M, Tsuboi H, Hatakeyama N, Endou A, Takaba H, Kubo M, Fabrice D, Minfray C, Joly-Pottuz L, Martin J M., Miyamoto A. A computational chemistry study on friction of h-MoS_2. Part I. Mechanism of single sheet lubrication. *J Phys Chem B* **113**: 16526–16536 (2009)

[6] Onodera T, Morita Y, Nagumo R, Miura R, Suzuki A, Tsuboi H, Hatakeyama N, Endou A, Takaba H, Dassenoy F, Minfray C, Joly-Pottuz L, Kubo M, Martin J M, Miyamoto A. A computational chemistry study on friction of h-MoS_2. Part II. Friction anisotropy. *J Phys Chem B* **114**: 15832–15838 (2010)

[7] Wang C Q, Li H S, Zhang Y S, Sun Q, Jia Y. Effect of strain on atomic-scale friction in layered MoS_2. *Tribol Int* **77**: 211–217(2014)

[8] Levita G, Cavaleiro A, Molinari E, Polcar T, Righi MC. Sliding properties of MoS_2 layers: Load and interlayer orientation effects. *J Phys Chem C* **118**: 13809–13816 (2014)

[9] Cahangirov S, Ataca C, Topsakal M, Sahin H, Ciraci S. Frictional figures of merit for single layered nanostructures. *Phys Rev Lett* **108**: 126103(2012)

[10] Prandtl L, Angew Z. A conceptual model to the kinetic theory of solid bodies. *Math Mech* **858**: 1–19 (1928)

[11] Tomlinson G A. A molecular theory of friction. *Philos Mag J Sci* **7**: 905–939 (1929)

[12] Mak K F, Lee C, Hone J, Shan J, Heinz T F, Atomically thin MoS_2: A new direct-gap semiconductor. *Phys Rev Lett* **105**: 136805 (2010)

[13] Wang Z, Zhao K, Li H, Liu Z, Shi Z, Lu J, Suenaga K, Joung S K, Okazaki T, Jin Z, Gu Z. Gao Z, Iijima S. Ultra-narrow WS_2 nanoribbons encapsulated in carbon nanotubes. *J Mater Chem* **21**: 171–180 (2011)

[14] Ataca C, Şahin H, Ciraci S. Stable single-layer MX_2 transition-metal oxides and dichalcogenides in a honeycomb-like structure. *J Phys Chem C* **116**: 8983–8999 (2012)

[15] Winkler B, Pickard C J, Segall M D, Milman V. Density functional study of charge ordering in $Cs_2Au(I)Au(III)Cl_6$ under pressure. *Phys Rev B* **63**: 14103 (2001)

[16] Wu Z, Cohen R E. More accurate generalized gradient approximation for solids. *Phys Rev B* **73**: 235116 (2006)

[17] Ceperley D M, Alder B J. Ground state of the electron gas by a stochastic method. *Phys Rev Lett* **45**: 566–569 (1980)

[18] Perdew J P, Zunger A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys Rev B* **23**: 5048–5079 (1981)

[19] Zhong W, Tomanek D. First-principles theory of atomic-scale friction. *Phys Rev Lett* **64**: 3054–3057 (1990)

[20] Girifalco L A, Hodak M. Van der Waals binding energies in graphitic structures. *Phys Rev B* **65**: 125404 (2002)

[21] Grimme S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J Comput Chem* **27**: 1787 (2006)

[22] Wu Z J, Zhao E J, Xiang H P, Hao X F, Liu X J, Meng J. Crystal structures and elastic properties of superhard IrN_2 and IrN_3 from first principles. *Phy Rev B* **76**: 054115 (2007)

[23] Rappe A K, Colwell K S, Casewit C J. Application of a universal force field to metal complexes. *Inorg Chem* **32**: 3438–3450 (1993)

[24] Dag S, Ciraci S. Atomic scale study of superlow friction between hydrogenated diamond surfaces. *Phys Rev B* **70**: 241401 (2004)

Jingyan NIAN. He received his master degree in physical chemistry in 2012 from Northwest Normal University, Lanzhou, China. During his undergraduate, he joined the State Key Laboratory of Solid Lubrication at Lanzhou Institute of Chemical Physics, Chinese Academy of Sciences. His current position is an associate researcher. His research area covers the atomic-scale friction, superlubricity, and space lubrication.

Liwei CHEN. He joined Prof. Zhiguang GUO's biomimetic materials of tribology (BMT) group at University of Hubei in 2014 in pursuing his Ph.D. degree. His current scientific interests are devoted to designing and fabricating superamphiphobic materials for the separation of oil/water mixtures.

Zhiguang GUO. He received his Ph.D. degree from Lanzhou Institute of Chemical Physics (LICP), Chinese Academy of Sciences (CAS) in 2007. During Feb. 2009 to Feb. 2010, he worked in Department of Physics, University of Oxford, UK, as a visiting scholar. Now he is a full professor in LICP financed by "Top Hundred Talents" program of CAS. Till now, he has published more than 130 papers about the interfaces of materials.

Weimin LIU. He received his Ph.D. degree with major of Lubricating Materials and Tribology from Lanzhou Institute of Chemical Physics (LICP), Chinese Academy of Sciences in 1990. In 2013, he was elected as the fellow of Chinese Academy of Sciences, and now, he is the director of the state key lab of solid lubrication. Up to now, he has published more than 500 papers and his research interests mainly focus on space lubrication and high performance lubricants.

Friction **5**(1): 32–44 (2017)
DOI 10.1007/s40544-016-0129-3
ISSN 2223-7690
CN 10-1237/TH
RESEARCH ARTICLE

Truncated separation method for characterizing and reconstructing bi-Gaussian stratified surfaces

Songtao HU[1], Weifeng HUANG[1], Noel BRUNETIERE[2], Xiangfeng LIU[1,*], Yuming WANG[1]

[1] *State Key Laboratory of Tribology, Tsinghua University, Beijing 100084, China*

[2] *Institut Pprime, CNRS-Universite de Poitiers-ENSMA, Futuroscope Chasseneuil Cedex 86962, France*

Received: 13 July 2016 / Revised: 21 September 2016 / Accepted: 03 November 2016

Abstract: Existing ISO segmented and continuous separation methods for differentiating the two components contained within a bi-Gaussian stratified surface were developed based on the fit of the probability material ratio curve. In the present study, because of the significant effect of the plateau component on tribological behavior such as asperity contact, wear and friction, a truncated separation method is proposed based on the truncation of the upper Gaussian component defined by zero skewness. The three separation methods are applied to real worn surfaces. Surface-separation and surface-reconstruction results show that the truncated method accurately captures the upper component identically to the ISO and continuous ones. The identification of the lower component characteristics requires performing a curve fit procedure on the data left after truncation. However, the truncated method fails in identifying the upper component when the material ratio of the transition is less than 9%.

Keywords: surface simulation; worn surface; stratified surface; mechanical face seal

1 Introduction

Surface texture can be considered as the fingerprint of a component [1, 2]. It can: (1) give an evaluation of the quality of manufacture process; (2) guide "reverse engineering" by revealing the unknown manufacture processes; (3) render the current state of the component itself; and (4) control a component's functional performance with respect to lubrication, asperity contact, wear, friction, etc. It is therefore imperative to research surface analysis methods. In addition to surface analysis, surface simulation is required because experiments usually require considerable financial inputs. Surface generation can also be used to vary surface statistical properties during tribological simulations.

To analyze rough surfaces, the central moment parameter set Rq (= σ), Rsk and RKu is widely used, and can well characterize the vast majority of rough surfaces [3]. Yet, it fails in assessing bi-Gaussian stratified surfaces such as two-process surfaces [4–6] and worn surfaces [2, 7–10]. The cylinder liner manufactured by plateau honing is a representative two-process surface which consists of smooth wear-resistant and load-bearing plateaus with intersecting deep valleys working as oil reservoirs and debris traps. During the wear process, a prepared rough surface is altered, yielding a two-scale rough surface with a large-scale roughness in the valleys and a small-scale roughness in the plateaus left by a truncation of the peaks of the large-scale roughness. Two main methods based on the material ratio curve (i.e., Abbott curve) [11] have been used to identify bi-Gaussian surfaces up to now: the Rk parameter set [12] and the probability

* Corresponding author: Xiangfeng LIU, E-mail: liuxf@tsinghua.edu.cn

material ratio method [13]. The Rk parameter set is illustrated in Fig. 1(a) referring to ISO 13565-2 [12]. It divides the material ratio curve into three regions, i.e., core, peak and valley regions, which conflicts with a two-stage manufacturing process. Thus, the probability material ratio curve method offers an effective choice. Williamson [14] found that the material ratio curve of a Gaussian surface plotted on a Gaussian standard deviation scale is linear; the intercept provides the mean value of asperity height whilst the slope gives Rq (= σ). A bi-Gaussian surface $z_{i,j}$ should thus embody two lines, as shown in Fig. 1(b) [9] referring to ISO 13565-3 [13]. Rpq (= σ_u) corresponds to the mean square root of the upper surface $z_{u,i,j}$, and Rvq (= σ_l) corresponds to the mean square root of the lower surface $z_{l,i,j}$. The knee-point (Rmq, z_k) defines the separation of the upper and the lower surfaces, and Pd [8] provides the distance between their mean planes (z_{mu}, z_{ml}). Whitehouse [4], Malburg et al. [5], Sannareddy et al. [6], Leefe [7], and Pawlus and Grabon [8] have used the probability material ratio curve method to analyze bi-Gaussian surfaces. However, all of these previous works were proceeded by using the segmented separation method to obtain the two components from a stratified surface. Hu et al. [2], recently, have criticized the two drawbacks of the segmented separation method. The first is that the segmented separation method arbitrarily assumes that the probability material ratio curve simply consists of two straight lines connected at a knee-point. In fact, the probability material ratio curve should exhibit a smooth transition region resulting from the gaps in the original plateau profile induced by the deep valleys in the rough profile [6] and the unity-area demand of the probability density function. The second drawback is that the segmented separation method tends to induce a large error in the small-scale roughness component. Therefore, Hu et al. [2] proposed a continuous separation method, which has been validated to perfectly overcome the drawbacks on simulated pure bi-Gaussian surfaces and real worn surfaces. Additionally, the continuous form of the probability density function in Ref. [2] was applied to modify the segmented stratified asperity contact model of Leefe [10]. However, when initially discussing the efficiency of the

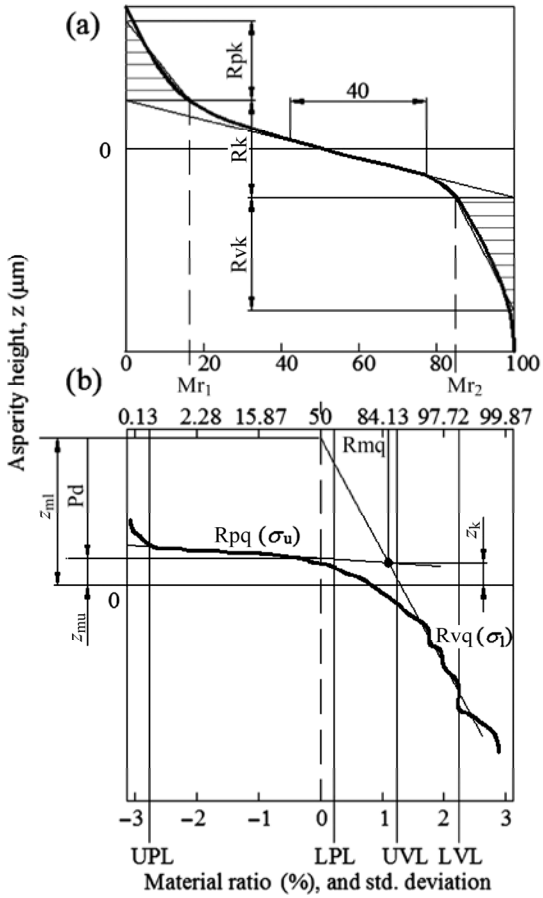


Fig. 1 Characterization of a two-process or worn surface [9].

continuous separation method in Ref. [2], Hu et al. selected a simple segmented method [6] rather than the procedure in ISO 13565-3 [13]. When they extended the application of the continuous method to the fields of lubrication and asperity contact [9], they compared the continuous and the ISO segmented separation methods.

To generate Gaussian distributed rough surfaces, three main models (i.e., autoregressive model [15–17], moving average model [18–20], and function series [21–23]) have been proposed together with either a direct approach or a fast Fourier transform (FFT) approach. To further generate a non-Gaussian surface, the Johnson translation system [24] together with auxiliary algorithms [25, 26] can be used to impose the target skewness and kurtosis values. Recently, Francisco and Brunetiere [27] developed an improved curve translation system. However, even if a translation system could help to reproduce a non-Gaussian

surface that captures the roughness, correlation lengths, skewness and kurtosis of a worn surface, it still loses the stratified characteristic. To address this limitation, Pawlus [28] specified the surface parameters for the two components, and generated a bi-Gaussian surface based on the superposition principle. Hu et al. [2, 9] also used the superposition principle to numerically generate bi-Gaussian surfaces. Furthermore, they combined the superposition principle (i.e., surface generation) with the probability material ratio curve method (i.e., surface analysis) to reconstruct real worn surfaces.

As we know, the upper component has a significant effect on the tribological behavior of a rough surface [9]. For instance, asperity contact usually appears in the peak region rather than the valley region. An error in modeling the asperity contact will induce a sequent deviation when analyzing wear or friction. In the present study, a truncated separation method is proposed based on the truncation of the upper Gaussian component, allowing capturing all surface peaks. Contrary to the ISO segmented and the continuous separation methods using the curve fit of the probability material ratio curve to separate the two Gaussian components, the truncated one adopts a sufficient condition (i.e., finding a threshold plane with the property that the surface defined by the points above the plane has zero skewness) to isolate the upper Gaussian surface. These three separation methods are applied to four real worn surfaces from Ref. [2]. A comparison is carried out on both surface separation and surface reconstruction. The limitation analysis of the new truncated method is also performed.

2 Existing approaches for separating and reconstructing worn surfaces

2.1 Surface separation method

In ISO 13565-3 [13], a segmented linear regression is proceeded after excluding the nonlinear regions. Its main procedure is as follows [9, 13]: (1) A conic section is used to initially fit the probability material ratio curve. Based on the conic-section asymptotes, a bisector line can be determined, and the transition point from plateau to valley is initially estimated by the intersection of the bisector line with the conic section. (2) The second derivative of the probability material ratio curve is calculated. A search, starting from the transition point, works upward through the plateau region and downward through the valley region. The upper limit of the plateau region UPL and the lower limit of the valley region LVL are assigned when the second derivative of the next point significantly exceeds the distribution of second derivatives of previous points on that side of the transition point. (3) The asperity-height axis of the probability material ratio curve is normalized with respect to the standard-deviation axis to yield a square region between UPL and LVL. Within the resulting square region, a conic-section fit is again used. Based on the corresponding asymptotes and three-time bisector line, the lower limit of the plateau region LPL and the upper limit of the valley region UVL are determined. (4) A segmented linear regression is applied to the non-normalized probability material ratio curve within [UPL, LPL] and [UVL, LVL], respectively, yielding component surface parameters σ_u, σ_l, z_{mu}, z_{ml} and Rmq.

In the continuous separation method, the continuous form of the material ratio curve is expressed as [2, 9, 10]

$$\mathrm{Mr} = 0.25 - 0.25\mathrm{erf}\left(\frac{z - z_{mu}}{\sqrt{2}\sigma_u}\right) - 0.25\mathrm{erf}\left(\frac{z - z_{ml}}{\sqrt{2}\sigma_l}\right) + 0.25\mathrm{erf}\left(\frac{z - z_{mu}}{\sqrt{2}\sigma_u}\right)\mathrm{erf}\left(\frac{z - z_{ml}}{\sqrt{2}\sigma_l}\right) \quad (1)$$

Then, the probability material ratio curve, G.S.D., is continuously rendered as [2, 9, 10]

$$\mathrm{G.S.D.} = \sqrt{2}\mathrm{inverf}\left[-0.5 - 0.5\mathrm{erf}\left(\frac{z - z_{mu}}{\sqrt{2}\sigma_u}\right) - 0.5\mathrm{erf}\left(\frac{z - z_{ml}}{\sqrt{2}\sigma_l}\right) + 0.5\mathrm{erf}\left(\frac{z - z_{mu}}{\sqrt{2}\sigma_u}\right)\mathrm{erf}\left(\frac{z - z_{ml}}{\sqrt{2}\sigma_l}\right)\right] \quad (2)$$

By applying Eq. (2) to fit the probability material ratio curve of the target surface, the values of z_{mu}, σ_u, z_{ml} and σ_l are determined. In order to calculate other surface parameters such as correlation length, summit density and mean summit curvature radius for each component, the knee-point concept from the segmented separation

method is retained. The knee-point (Rmq, z_k) can be deduced from two points (0, z_{mu}) and (0, z_{ml}) [2, 9, 10]

$$\text{Rmq (in std. deviation scale)} = \frac{z_{mu} - z_{ml}}{\sigma_u - \sigma_l} \tag{3a}$$

$$z_k = \frac{z_{ml}\sigma_u - z_{mu}\sigma_l}{\sigma_u - \sigma_l} \tag{3b}$$

2.2 Surface reconstruction approach

The kernel of the bi-Gaussian reconstruction approach is to generate two individual Gaussian surfaces based on the results of surface separation, and then to combine the two Gaussian components according to the superposition principle. Therefore, the first step of the bi-Gaussian reconstruction approach is to generate Gaussian surfaces. To generate a Gaussian surface with specified σ and correlation lengths, asperity height $z_{i,j}$ in the moving average model is expressed as [2, 3, 9, 18–20]

$$z_{i,j} = \sum_{k=-N/2}^{N/2} \sum_{l=-M/2}^{M/2} a_{k,l}\eta_{i+k,j+l} \tag{4}$$

η is the Gaussian random series. $a_{k,l}$ is a matrix of coefficients that can be obtained from the autocorrelation function ACF [18]

$$\text{ACF}(p,q) = \sum_{k=-N/2+p}^{N/2} \sum_{l=-M/2+q}^{M/2} a_{k,l}a_{k-p,l-q} \tag{5}$$

To avoid a complex solution of the nonlinear system of (M+1)×(N+1), the FFT is used. Equation 5 is transformed into the frequency domain [20]

$$\text{PSD}(\omega_p, \omega_q) = H(\omega_p, \omega_q) \times \bar{H}(\omega_p, \omega_q) \tag{6}$$

where PSD (power spectral density) is the FFT of the ACF, and H is the FFT of $a_{k,l}$. As the PSD is composed only of real coefficients, Eq. (6) can be rewritten as

$$H(\omega_p, \omega_q) = \sqrt{\text{PSD}(\omega_p, \omega_q)} \tag{7}$$

$z_{i,j}$ can be calculated from Eq. (4) after applying an inverse FFT to Eq. (7) in order to obtain $a_{k,l}$. $z_{i,j}$ can also be obtained by applying an inverse FFT to the scalar product between the FFT of η and Eq. (7) [20].

According to Refs. [2, 9, 28], with the specified component surface parameters σ_u, σ_l and autocorrelation functions, two Gaussian component surfaces $z_{u,i,j}$ and $z_{l,i,j}$ are numerically generated, respectively. The distance Pd between the mean planes z_{mu} and z_{ml} is calculated by Pd = z_{mu}–z_{ml}. The target rough surface $z_{i,j}$ can then be reproduced by using the superposition principle, i.e., $z_{i,j}$ = min ($z_{u,i,j}$, $z_{l,i,j}$). The newly combined surface $z_{i,j}$ is then updated by moving it to its mean plane.

3 Truncated separation method

Figure 2 shows a schematic diagram of the truncated separation method. For a worn surface, its upper Gaussian component can be defined by zero skewness and a certain value of the kurtosis close to three [29]. Thus, a sufficient condition to isolate this component is to find a threshold plane with the property that the surface defined by the points above the plane has zero skewness. By this, the remaining points (below the threshold plane) are grouped into the lower component. The reasons to select the skewness as the convergence index are: (1) The kurtosis is a quartic central moment and is much more sensitive than the skewness to the selection of the threshold plane. (2) The skewness well indicates the asymmetry of the surface. For a Gaussian surface which has a symmetric shape relative to its mean plane, the skewness is zero. For an asymmetrically distributed surface, the skewness will be negative if the surface has more peaks under the mean plane, while a positive skewness is specific to a surface with high peaks and shallow valleys. As a result, it is easy to apply a dichotomy method to the skewness when searching the threshold plane. The basic idea of the dichotomy method is described as follows: (1) For the target worn surface $z_{i,j}$, boundaries are defined as z_{start} = min($z_{i,j}$) and z_{end} = max($z_{i,j}$). Since the skewness is a dimensionless number and is expected to be zero, the convergence condition is whether the absolute value of the skewness is smaller than a specified coef_{Rsk}. (2) The threshold plane is set as z_{k0} = (z_{start}+z_{end})/2. (3) Based on z_{k0}, $z_{i,j}$ is divided into two components. The skewness of the upper component is calculated. (4) If the absolute value of the calculated skewness is smaller than coef_{Rsk}, the iterative procedure is completed, and the required threshold

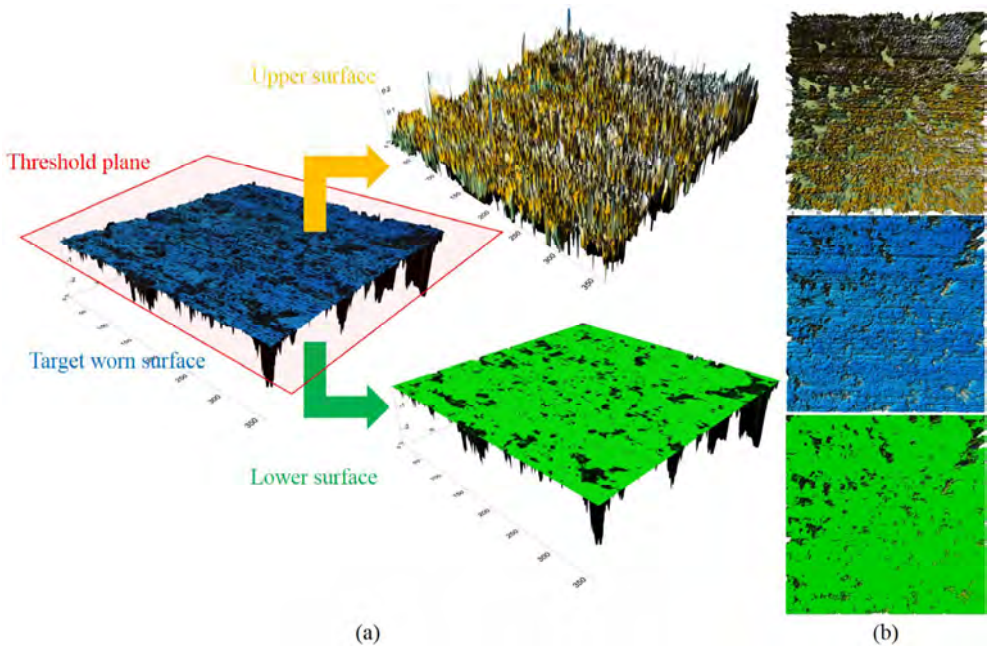


Fig. 2 Schematic of the truncated separation method: (a) 3D and (b) 2D.

plane is z_{k0}. Otherwise, z_{start} (or z_{end}) is set to z_{k0} if the skewness is negative (or positive). (5) Repeating steps 2–4 until the iterative procedure satisfies the convergence condition.

After the upper component is separated from the target worn surface, the lower component is retained. Logically, the surface parameters of the lower component can be simply calculated referring to these remaining data. It is named Route A. Unfortunately, the Rmq of a real worn surface is usually larger than 50%, indicating an abundant loss of points in the remaining lower component, as shown in Fig. 2. By this, the left information is not sufficient to accurately calculate σ_l and z_{ml}. In such a case, the linear regression of the probability material ratio curve is still required to the left points, yielding σ_l (slope) and z_{ml} (intercept). It is named Route B.

4 Results and discussion

4.1 Results

Four real worn seal surfaces from Refs. [2, 10] are used. These four surfaces from a mono-spring face seal (Fig. 3) have undergone a running-in period of 24 hours and an additional 100 hours of service. They are, respectively, made of silicon carbide (SiC), tungsten carbide (TC), resin-impregnated carbon (RiC) and metal-impregnated carbon (MiC). A white light interferometer Talysurf CCI with a 50× objective is used to measure the above surfaces. Because the interferometer is unable to receive some signals when the local slope of the surface is greater than 27°, inducing noises at the boundary of deep valleys, a median filter is applied for noise reduction. The unmeasured points are retained without any interpolation.

In line with previous works [2, 3, 9, 10], the fundamental surface parameters are calculated as

Fig. 3 Mono-spring face seal [2, 9, 10].

$$\sigma = \sqrt{\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M} z_{i,j}^2} \tag{8a}$$

$$\mathrm{Rsk} = \frac{1}{\sigma^3}\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M} z_{i,j}^3 \tag{8b}$$

$$\mathrm{Rku} = \frac{1}{\sigma^4}\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M} z_{i,j}^4 \tag{8c}$$

$$\mathrm{ACF}(p,q) = \frac{1}{\sigma^2}\frac{1}{NM}\sum_{i=1}^{N-p}\sum_{j=1}^{M-q} z_{i,j} z_{i+p,j+q} \tag{9}$$

$$\mathrm{ACF}\left(\frac{\lambda_x}{\Delta_x}, 0\right) = 0.2 \tag{10a}$$

$$\mathrm{ACF}\left(\frac{\lambda_y}{\Delta_y}, 0\right) = 0.2 \tag{10b}$$

M and N are the numbers of points in the x and y directions, both equaling to 1,024; λ_x and λ_y are the 80% correlation lengths in the x and y directions; Δ_x and Δ_y are the x- and y-direction sampling intervals. Since the size of each surface is 360 μm × 360 μm, Δ_x and Δ_y are both equal to 0.352 μm. The results of the surface-parameter calculation are listed in Table 1. Note that in Refs. [2, 9], two mistakes were made. (1) Variables p and q in Eq. (9) were limited to 30 which was narrow compared to the real correlation lengths of the RiC and the MiC surfaces. Thus, in Refs. [2, 9], the resulting correlation lengths of the RiC and the MiC surfaces were 10.6 μm. (2) λ_x and λ_y were switched. These two mistakes did not affect the conclusions.

The truncated, the continuous and the ISO segmented separation methods are applied to the above target surfaces, successively. Firstly, Route A is selected to deal with the lower components. The results are displayed in the left part of Fig. 4 and are listed in

Table 1 Surface parameters of the target surfaces.

Parameter	Value			
	SiC	TC	RiC	MiC
σ (μm)	0.0922	0.0339	0.0641	0.200
Rsk	−1.56	−10.0	−4.00	−6.50
Rku	7.46	146	34.3	59.9
λ_x (μm)	4.25	4.25	15.6	12.4
λ_y (μm)	3.90	4.25	9.22	11.0
Ratio of unmeasured points (%)	0.133	0.0309	0.245	1.03

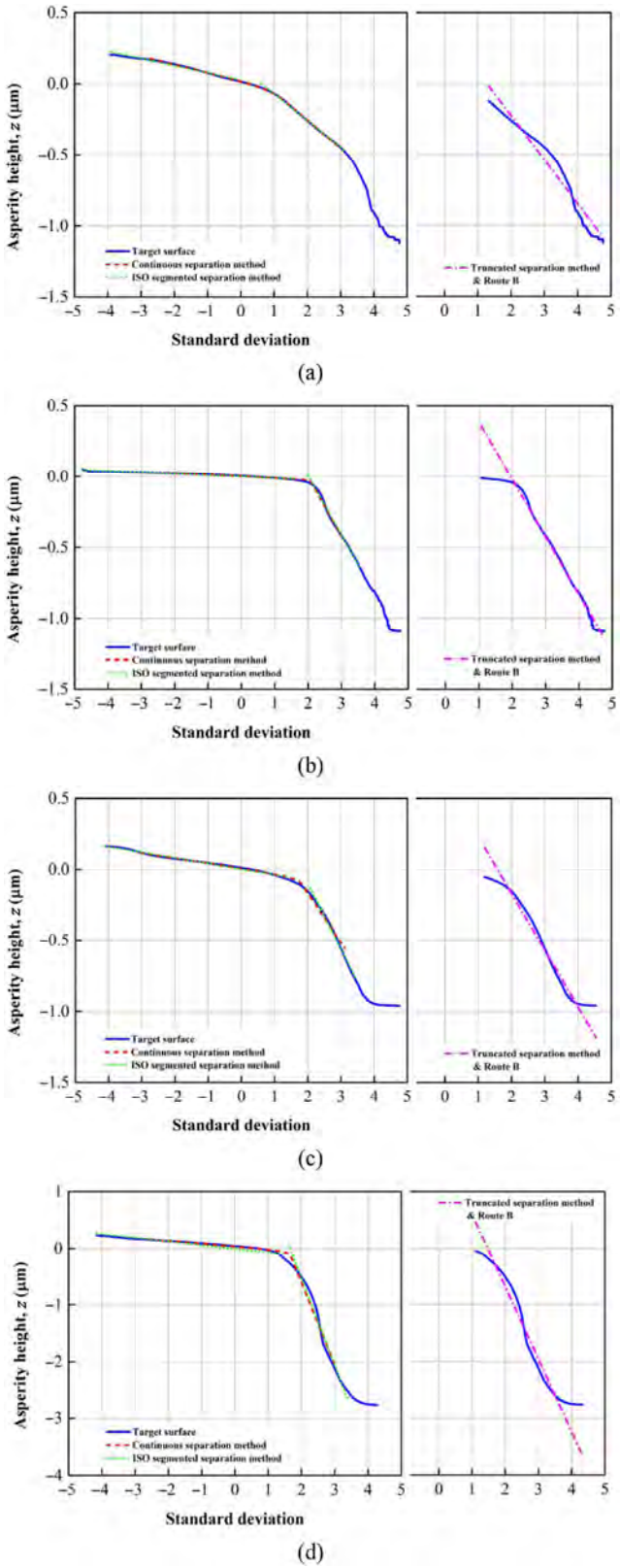


Fig. 4 Component separation by using the continuous, the ISO (left part) segmented and the truncated with Route B (right part) separation methods: (a) SiC, (b) TC, (c) RiC, and (d) MiC.

Table 2. Note that the continuous and the ISO segmented methods are carried out within [–3, 3] [2, 9, 10] because this interval has covered the vast majority of the material ratio curve from 0.13% to 99.87%. In the use of the truncated separation method, $coef_{Rsk}$ is set to 1×10^{-5}.

It can be concluded from Fig. 4 and Table 2 that [9]: both the ISO segmented and the continuous separation methods can well separate the components from the target surface; the continuous one yields results almost identical to the ISO one but in a much more efficient way. In terms of the present truncated separation method with Route A, it well separates the upper component. However, it exhibits two drawbacks. (1) The resulting lower component is obviously different from those obtained by the other two methods. As mentioned in Section 3, the deviation should be caused by the abundant loss of points in the lower part when Rmq is much larger than 50%. (2) There is a significant deviation between the truncated method and the other two in the knee-point. In the truncated separation method, the threshold plane z_k is firstly determined by the dichotomy method mentioned in Section 3. Then, Rmq can be calculated based on the known z_k. The resulting knee-point is thus a certain point in the probability material ratio curve. However, in the ISO segmented and the continuous separation methods, σ_u, σ_l, z_{mu} and z_{ml} are firstly determined. Then, Rmq and z_k are calculated according to Eq. (3). The knee-point usually does not really locate on the probability material curve.

To solve the first drawback, Route B is used as shown in the right part of Fig. 4, thus yielding Table 3. In Table 3, the quality of the lower component has been greatly improved especially for the latter three cases. In contrast with a simulated bi-Gaussian case whose two components are known beforehand, it is hard to assess which way used for calculating the knee-point is better with respect to real worn surfaces. Thus, this topic will be discussed in Section 4.2.

Using the surface-separation results from Table 3, surface reconstructions, based on the approach presented in Section 2.2, are performed. Two tips should be emphasized. (1) For a bi-Gaussian reconstruction approach, it is difficult to select the input correlation lengths for a component. According to Refs. [2, 9], the input two correlation lengths of each component in the present study are arbitrarily set to the same value as the whole target surface shown in Table 1. (2) During the Gaussian surface generation, the ACF is assumed to be an exponential function. The surface parameters of the reconstructed surfaces are calculated and listed in Table 4. The comparison is also conducted in terms of shaded relief map (Fig. 5) and probability material ratio curve (Fig. 6). It can be seen from Table 4 that the truncated reconstructed surface exhibits surface properties almost identically to the ISO and the continuous reconstructed ones especially in the latter

Table 2 Surface parameters of the components using the truncated, the continuous and the ISO segmented separation methods where Route A is used.

		σ_u (μm)	σ_l (μm)	z_{mu} (μm)	z_{ml} (μm)	Knee-point (Rmq, z_k) (%, μm)	Rsk_u ($\times 10^{-6}$)	Rku_u
SiC	Trun.&A	0.0585	0.0834	0.0220	–0.215	(90.7, –0.121)	3.97	2.85
	Cont.	0.0548	0.198	0.0419	0.142	(75.7, 0.00370)		
	ISO	0.0503	0.189	0.0271	0.121	(75.0, –0.00685)		
TC	Trun.&A	0.00894	0.0744	0.00694	–0.0417	(85.8, –0.0113)	2.86	2.30
	Cont.	0.0115	0.425	0.00253	0.851	(98.0, –0.0211)		
	ISO	0.00974	0.417	0.000598	0.839	(98.0, –0.0194)		
RiC	Trun.&A	0.0312	0.103	0.0163	–0.123	(88.3, –0.0531)	2.88	2.84
	Cont.	0.0353	0.351	0.00898	0.536	(95.2, –0.0499)		
	ISO	0.0397	0.468	0.00185	0.841	(97.5, –0.0758)		
MiC	Trun.&A	0.0419	0.406	0.0505	–0.329	(86.7, –0.0487)	–1.34	2.70
	Cont.	0.0488	1.45	0.0379	2.29	(94.6, –0.0407)		
	ISO	0.0651	1.54	–0.00709	2.56	(95.9, –0.120)		

Table 3 Surface parameters of the components using the truncated, the continuous and the ISO segmented separation methods where Route B is used.

		σ_u (μm)	σ_l (μm)	z_{mu} (μm)	z_{ml} (μm)	Knee-point (Rmq, z_k) (%, μm)	Rsk_u ($\times10^{-6}$)	Rku_u
SiC	Trun.&B	0.0585	0.311	0.0220	0.397	(90.7, –0.121)	3.97	2.85
	Cont.	0.0548	0.198	0.0419	0.142	(75.7, 0.00370)		
	ISO	0.0503	0.189	0.0271	0.121	(75.0, –0.00685)		
TC	Trun.&B	0.00894	0.405	0.00694	0.791	(85.8, –0.0113)	2.86	2.30
	Cont.	0.0115	0.425	0.00253	0.851	(98.0, –0.0211)		
	ISO	0.00974	0.417	0.000598	0.839	(98.0, –0.0194)		
RiC	Trun.&B	0.0312	0.400	0.0163	0.630	(88.3, –0.0531)	2.88	2.84
	Cont.	0.0353	0.351	0.00898	0.536	(95.2, –0.0499)		
	ISO	0.0397	0.468	0.00185	0.841	(97.5, –0.0758)		
MiC	Trun.&B	0.0419	1.28	0.0505	1.89	(86.7, –0.0487)	–1.34	2.70
	Cont.	0.0488	1.45	0.0379	2.29	(94.6, –0.0407)		
	ISO	0.0651	1.54	–0.00709	2.56	(95.9, –0.120)		

Table 4 Surface parameters of the target and the reconstructed surfaces.

		σ (μm)	Rsk	Rku
SiC	Target	0.0922	–1.56	7.46
	Trun.&B	0.0859	–2.09	12.8
	Cont.	0.0926	–1.51	6.84
	ISO	0.0882	–1.57	6.98
TC	Target	0.0339	–10.0	146
	Trun.&B	0.0364	–10.2	139
	Cont.	0.0358	–10.2	146
	ISO	0.0341	–10.7	158
RiC	Target	0.0641	–4.00	34.3
	Trun.	0.0705	–5.03	40.5
	Cont.	0.0674	–4.19	31.8
	ISO	0.0675	–4.72	44.1
MiC	Target	0.200	–6.50	59.9
	Trun.	0.227	–5.95	46.6
	Cont.	0.224	–6.56	56.8
	ISO	0.214	–6.79	63.9

three cases. Yet, in the SiC case, the truncated separation method leads to an obvious deviation from the measured one. In Fig. 6, the upper part of the truncated reconstructed surface is still in a great agreement with that of the measured surface. In other words, the above error should arise from the lower part. In fact, as shown in Figs. 4 and 6, the measured SiC surface has a lower component with small slope, and the nonlinear region caused by deep scratches or outlying valleys occupies a considerable proportion in the lower component. In such a case, because all data in the lower component are utilized, the truncated separation method will surely yield a solution to the lower component different from the other two methods.

4.2 Discussion

For the truncated separation method, Rmq is of huge importance with respect to the accuracy. If Rmq locates in the nonlinear regions at the two ends of the probability material ratio curve, it corresponds to a single-stratum surface beyond the bi-Gaussian scope. In such a case, both the ISO segmented and the continuous separation methods lose the efficiency. If Rmq locates in the linear or the transition region, in principle, the truncated separation method can always find a threshold plane to obtain a Gaussian surface, and the accuracy of the upper component identification can be guaranteed. For the accuracy of the lower component characterization, logically, the truncated separation method with Route B can satisfy the full bi-Gaussian scope. The truncated separation method with Route A may only suit for the situation where Rmq is much smaller than 50%.

Because Rmq of the above real worn surfaces are all much larger than 50%, and because the exact information of the components is unknown, surface simulation is used to gauge the impact of Rmq on the accuracy of the truncated method. A series of

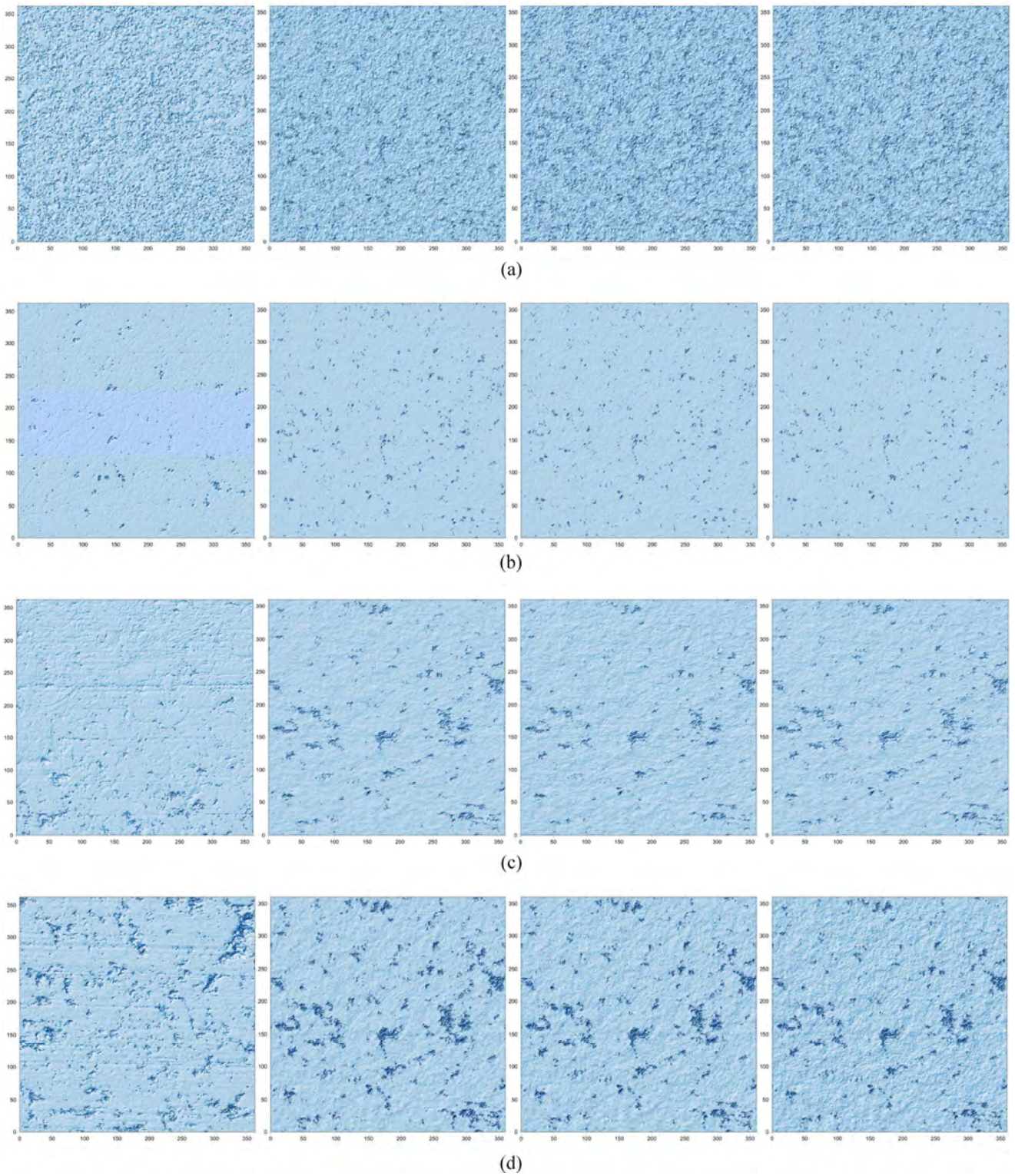


Fig. 5 Visualization of the target and the reconstructed surfaces (from left to right: target, truncated bi-Gaussian reconstructed, continuous bi-Gaussian reconstructed, and ISO segmented bi-Gaussian reconstructed). (a) SiC; (b) TC; (c) RiC; (d) MiC.

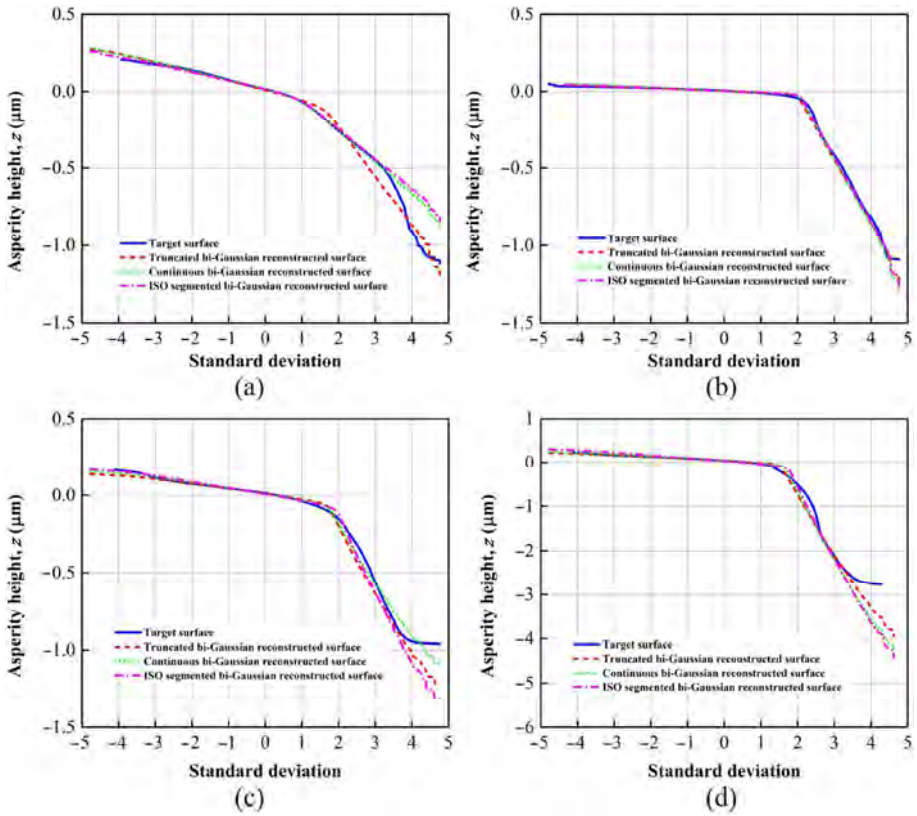


Fig. 6 Probability material ratio curves of the target and the reconstructed surfaces: (a) SiC; (b) TC; (c) RiC; (d) MiC.

bi-Gaussian surfaces is generated. Referring to the component surface parameters in Table 3, components are set as $\sigma_u = 0.05$ μm, $\sigma_l = 0.5$ μm, $z_{ml} = 0$ μm, and $\lambda_{xu} = \lambda_{yu} = \lambda_{xl} = \lambda_{yl} = 10.2$ μm. According to Eq. (3a), Pd is, respectively, set to 0.740, 0.700, 0.664, 0.632, 0.603, 0.577 0.466, 0.379, 0.304, 0.236, 0.173, 0.114, 0.0565 and 0 μm, expecting to guarantee Rmq correspond to 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% and 50%, respectively. By using these specified components, ten bi-Gaussian surfaces are produced by using the surface reconstruction approach introduced in Section 2.2. Then, the truncated separation method is applied to these numerically generated bi-Gaussian surfaces with Routes A and B, respectively. The results are displayed in Fig. 7.

In Fig. 7(a), two stages can be obviously observed: stage 1 (the expected Rmq is lower than 9%) and stage 2 (the expected Rmq is greater than 9%). In

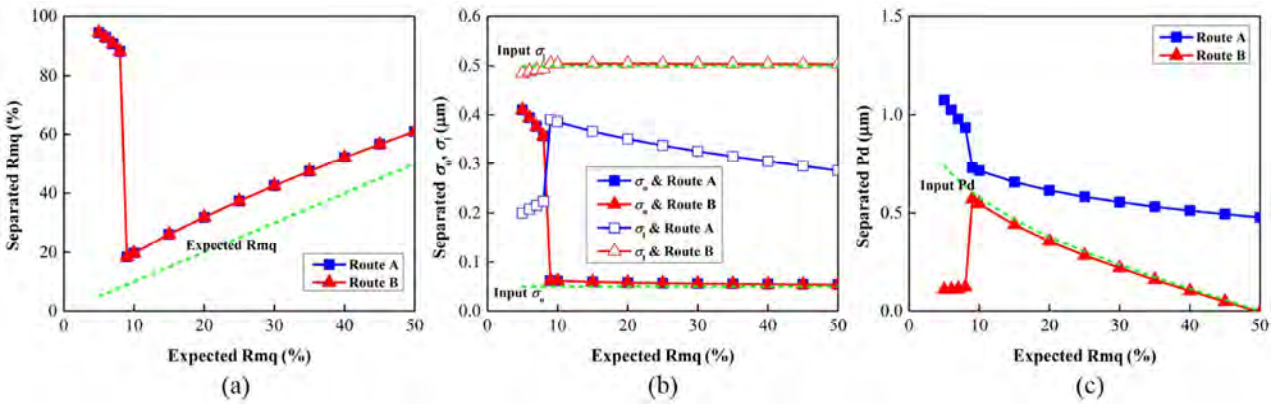


Fig. 7 Accuracy analysis of the truncated separation method with Routes A and B: (a) Rmq; (b) σ_u and σ_l; (c) Pd.

stage 1, the separated Rmq is much greater than the expected one. It is because that in such a situation, the input upper component nearly vanishes, and therefore the so-called "bi-Gaussian" surface behaves more like a single-stratum Gaussian surface. It can be validated by observing σ_u which is nearly 0.4 μm in Fig. 7(b). Namely, the truncated separation method isolates a surface with zero skewness that is almost equal to the input lower Gaussian component. In stage 2, the separated Rmq has a similar tendency with the expected one, but has a constant deviation. This deviation may arise from two sources: system error and method error. This topic will be further discussed in the following paragraph. With respect to Fig. 7(b), the truncated separation method with any route well captures the input upper component out of the influence of the expected Rmq in the bi-Gaussian scope (i.e., stage 2). For capturing the input lower component, only Route B gives good results. Although the accuracy of Route A increases with the decrease of the expected Rmq, σ_l obtained by Route A is still much smaller than 0.5 μm when the expected Rmq reaches the transition point between stage 1 and stage 2. In Fig. 7(c), the separated Pd by Route B is in a great agreement with the input value in stage 2, while exhibits a deviation in stage 1. For Route A, it renders a Pd value with an obvious deviation regardless of the expected Rmq.

As previously described, there is a deviation between the separated Rmq and the expected Rmq in the truncated separation method. As mentioned above, Pd is a direct variable in the surface reconstruction approach to indirectly set Rmq. Thus, a system error may be produced during surface generation. It is also possible that the truncated separation method itself has a defect in capturing Rmq. Therefore, the ISO segmented and the continuous separation methods are also applied to the ten simulated bi-Gaussian surfaces, yielding Fig. 8(a). It is obvious that the other two separation methods also capture Rmq with a deviation from the expected value. The tiny deviation in the continuous method reveals that the system error can be ignored. Thus, the deviation is due to the method error. As shown in Figs. 7(b) and 7(c), in stage 2, the truncated separation method with Route B has yielded σ_u, σ_l and Pd almost identically to the input ones.

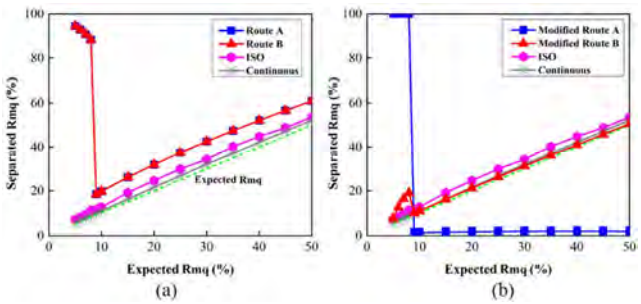


Fig. 8 Accuracy comparison between the truncated, the ISO segmented and the continuous separation methods with respect to *Rmq*: (a) real results; (b) modified results.

According to Eq. (3a), the resulting Rmq, logically, should be very close to those obtained in the ISO segmented and the continuous methods. Therefore, the great deviation in the truncated method with Route B should be caused by the different ways used for calculating the knee-point which has been mentioned in Section 4.1. To demonstrate this deduction, σ_u, σ_l and Pd from Figs. 7(b) and 7(c) are substituted into Eq. (3a) to calculate a modified Rmq value, yielding Fig. 8(b). It is obvious that in stage 2, the deviation in the truncated separation method with Route B becomes much smaller, and is even smaller than that in the continuous method.

It can be concluded that: (1) for the truncated separation method, the identification of the lower component characteristics needs a curve fit procedure on the data left after truncation; (2) the modified truncated separation method can guarantee the accuracy of the lower component, but fails in identifying the upper component when Rmq is too small (less than 9%); (3) the Rmq deviation between the modified truncated method and the other two methods arises from the different ways used for calculating the knee-point.

5 Conclusions

Existing ISO segmented and continuous separation methods for differentiating the two components contained within a bi-Gaussian stratified surface are established based on the fit of the probability material ratio curve. Because of the significant effect of the plateau component on tribological behavior such as asperity contact, wear and friction, a truncated separation method is proposed. This method is based on

the truncation of the upper Gaussian component defined by zero skewness. The new separation method together with the other two are applied to four real worn surfaces. The results of surface separation and surface reconstruction show that the truncated method can accurately capture the upper component identically to the ISO and the continuous methods. To accurately determine the statistical properties of the lower component, it is necessary to perform a curve fit procedure on the data left after truncation. Rmq is recommended to be calculated based on the determined statistical properties of the two components. The limitation of the truncated method is that it fails in identifying the upper component characteristics when Rmq is too small (less than 9%).

Acknowledgements

This work was supported by the National Key Basic Research (973) Program of China (No. 2012CB026003), the National Science and Technology Major Project (No. ZX06901), and the National Science and Technology Support Plan Projects (No. 2015BAA08B02).

References

[1] Whitehouse D J. Surfaces-a link between manufacture and function. *Proc Inst Mech Eng* **192**: 179–188 (1978)

[2] Hu S, Brunetiere N, Huang W, Liu X, Wang Y. Continuous separating method for characterizing and reconstructing bi-Gaussian stratified surfaces. *Tribol Int* **102**: 454–452 (2016)

[3] Minet C, Brunetiere N, Tournerie B, Fribourg D. Analysis and modeling of the topography of mechanical seal faces. *Tribol Trans* **53**: 799–815 (2010)

[4] Whitehouse D J. Assessment of surface finish profiles produced by multi-process manufacture. *Proc the Inst Mech Eng Part B: J Eng Manufact* **199**: 263–270 (1985)

[5] Malburg M C, Raja J, Whitehouse D J. Characterization of surface texture generated by plateau honing process. *CIRP Annals-Manufacturing Technology* **42**: 637–639 (1993)

[6] Sannareddy H, Raja J, Chen K. Characterization of surface texture generated by multi-process manufacture. *Int J Mach Tools Manufact* **38**: 529–536 (1998)

[7] Leefe S E. Bi-Gaussian' representation of worn surface topography in elastic contact problems. *Tribol Ser* **34**: 281–290 (1998)

[8] Pawlus P, Grabon W. The Method of Truncation parameters measurement from material ratio curve. *Prec Eng* **32**: 342–347 (2008)

[9] Hu S, Huang W, Brunetiere N, Song Z, Liu X, Wang Y. Stratified effect of continuous bi-Gaussian rough surface on lubrication and asperity contact. *Tribol Int* **104**: 328–341 (2016)

[10] Hu S, Brunetiere N, Huang W, Liu X, Wang Y. Stratified revised asperity contact model for worn surfaces. *J Tribol* in press, DOI 10.1115/1.4034531 (2016)

[11] Abbot E J, Firestone F A. Specifying surface quality. *Mech Eng* **55**: 569–578 (1933)

[12] Surface texture: Profile method; surfaces having stratified functional properties—Part 2: Height characterization using the linear material ratio curve. ISO 13565-2, 1996.

[13] Surface texture: Profile method; surfaces having stratified functional properties—Part 3: Height characterization using the material probability curve. ISO 13565-3, 1998.

[14] Williamson J P B. Microtopography of surfaces. *Proc Inst Mech Eng* **182**: 21–30 (1985)

[15] Staufert G. Characterization of random roughness profiles —A comparison of AR-modeling technique and profile description by means of commonly used parameters. *Annals of the CIRP* **28**: 431–435 (1979)

[16] DeVries W R. Autoregressive time series models for surface profile characterization. *Annals of the CIRP* **28**: 437–440 (1979)

[17] Whitehouse D J. The generation of two dimensional random surfaces having a specified function. *Annals of the CIRP* **32**: 495–498 (1983)

[18] Patir N. A Numerical method for random generation of rough surfaces. *Wear* **47**: 263–277 (1978)

[19] Bakolas V. Numerical generation of arbitrarily oriented non-Gaussian three-dimensional rough surfaces. *Wear* **254**: 546–554 (2004)

[20] Hu Y Z, Tonder K. Simulation of 3-D random rough surface by 2-D digital filter and Fourier analysis. *Int J Mach Tools Manufact* **32**: 83–90 (1992)

[21] Majumdar A, Tien C. Fractal characterization and simulation of rough surfaces. *Wear* **136**: 313–327 (1990)

[22] Wu J. Simulation of rough surfaces with FFT. *Tribol Int* **33**: 47–58 (2000)

[23] Wu J. Simulation of non-Gaussian surfaces with FFT. *Tribol Int* **37**: 339–346 (2004)

[24] Johnson N L. Systems of frequency curves generated by method of translation. *Biometrika* **36**: 149–176 (1949)

[25] Watson W, Spedding T A. The time series modelling of non-Gaussian engineering processes. *Wear* **83**: 215–231 (1982)

[26] Hill I D, Hill R, Holder R L. Fitting Johnson curves by moments. *Applied Statistics* **25**: 180–189 (1976)

[27] Francisco A, Brunetiere N. A Hybrid method for fast and efficient rough surface generation. *IMechE Part J: J Eng Tribol* **230**: 747–768 (2016)

[28] Pawlus P. Simulation of stratified surface topographies. *Wear* **264**: 457–463 (2008)

[29] Tomescu A. Simulation of surface roughness for tribological applications. Master thesis. Universite de Poitiers, Poitiers, France, 2012.

Songtao HU. He is a Ph.D. student since 2012 in Department of Mechanical Engineering, Tsinghua University, China. His research interest is mechanical face seals.

Weifeng HUANG. He is currently an associate professor in Department of Mechanical Engineering, Tsinghua University, China. His research interest is mechanical face seals.

Noel BRUNETIERE. He is currently a CRNS researcher in Institut Pprime, France. His research interest is tribology, lubrication and fluid sealing.

Xiangfeng LIU. He is currently a professor and Ph.D. candidate supervisor in Department of Mechanical Engineering, Tsinghua University, China. His research interest is machine design and mechanical face seals.

Yuming WANG. He is currently a professor and a Ph.D. candidate supervisor in Department of Mechanical Engineering, Tsinghua University, China, and an academician of Chinese Academy of Engineering. His research interest is the fluid sealing technology.

Friction **5**(1): 45–55 (2017)
DOI 10.1007/s40544-016-0136-4
ISSN 2223-7690
CN 10-1237/TH
RESEARCH ARTICLE

Reduction of friction by normal oscillations. I. Influence of contact stiffness

M. POPOV[1,2,3,*], **V. L. POPOV**[1,2,3], **N. V. POPOV**[1]

[1] *Berlin University of Technology, Berlin 10623, Germany*

[2] *Tomsk Polytechnic University, Tomsk 634050, Russia*

[3] *Tomsk State University, Tomsk 634050, Russia*

Received: 04 June 2016 / Revised: 25 September 2016 / Accepted: 29 November 2016

Abstract: The present paper is devoted to a theoretical analysis of sliding friction under the influence of oscillations perpendicular to the sliding plane. In contrast to previous works we analyze the influence of the stiffness of the tribological contact in detail and also consider the case of large oscillation amplitudes at which the contact is lost during a part of the oscillation period, so that the sample starts to "jump". It is shown that the macroscopic coefficient of friction is a function of only two dimensionless parameters—a dimensionless sliding velocity and dimensionless oscillation amplitude. This function in turn depends on the shape of the contacting bodies. In the present paper, analysis is carried out for two shapes: a flat cylindrical punch and a parabolic shape. Here we consider "stiff systems", where the contact stiffness is small compared with the stiffness of the system. The role of the system stiffness will be studied in more detail in a separate paper.

Keywords: sliding friction; out-of-plane oscillation; contact stiffness; coefficient of friction; active control of friction

1 Introduction

The influence of vibration on friction is of profound practical importance [1]. This phenomenon is used in wire drawing [2, 3], press forming [4] and many other technological applications. Experimental studies of the influence of ultrasonic oscillations on friction started in the late 1950s [5]. In the subsequent years several illuminating works were performed using various techniques, e.g., measurement of electrical conductivity of the contact [6, 7]. Reduced friction has been observed both with oscillations in the contact plane (in-plane) [8] and perpendicular to it (out-of-plane) [9]. In the 2000s, interest in the interaction of friction and oscillations was promoted by applications such as traveling wave motors [10, 11] and the rapidly developing field of nanotribology [12, 13]. In recent years, detailed studies of the influence of ultrasonic oscillations and comparisons with various theoretical models have been performed by Chowdhury et al. [14] and Popov et al. for in-plane oscillations [15], and by Teidelt et al. for out-of-plane oscillations [16]. The latter paper also includes a comprehensive overview of previous works in the field up to 2012.

The above works provided an empirical basis for a qualitative understanding of the influence of oscillations on friction. However, good quantitative correspondence between experimental results and theoretical models could never be achieved (see, e.g., a detailed discussion in Ref. [17]), so it is not clear whether we adequately understand the physics of this phenomenon. Even the question of which oscillation properties determine the reduction of friction force is still under discussion: While in the case of static friction it seems to be the amplitude of displacement oscillation [15], for sliding friction it is believed to be the amplitude of velocity oscillation [11]. In the following, we will show that, in general, friction under oscillation is determined by both of these parameters.

The main novelty of the present paper compared to

* Corresponding author: M. POPOV, E-mail: m@popov.name

earlier work on the influence of oscillation on friction is explicit consideration of the *contact stiffness*. The influence of the contact stiffness is closely related to a fundamental and still unresolved question about the physical nature of the characteristic length determining the crossover from static friction to sliding. In earlier works on this topic, it was assumed that this characteristic length is an intrinsic property of a frictional couple and that its physical nature is rooted in microscopic interactions between the surfaces [15]. However, later investigations suggested another interpretation. Studies of friction in stick-slip microdrives [17, 18] have shown that the static and dynamic behavior of drives can be completely understood and precisely described without any fitting parameters just by assuming that the characteristic length responsible for the "pre-slip" during tangential (in-plane) loading of a contact is equivalent to partial slip in a tangential contact of bodies with curved surfaces. This contact-mechanical approach was substantiated in Ref. [19] by a theoretical study of the influence of in-plane oscillations on the static force of friction. It was shown that the characteristic length is simply the indentation depth multiplied by the coefficient of friction. Later, it was found that this is valid independently of the shape of the contact and also holds true for rough surfaces [20]. This hypothesis of the purely contact mechanical nature of the pre-slip and of the characteristic amplitude was verified experimentally for a wide range of radii of curvature and applied forces in Refs. [21, 22]. It was thus confirmed that describing friction under oscillation, including pre-slip, is basically a matter of correct contact mechanics and that the main governing parameter for both normal and tangential oscillation is the indentation depth. This realization also led to new generalizations in the physics of friction [23, 24], which, however, still need experimental verification.

In the present paper we utilize this new understanding of the importance of the precise contact mechanics and the key property of contact stiffness when considering the details of frictional processes. We focus our attention on the influence of normal (out-of-plane) oscillations on the macroscopic frictional force. We begin by looking at a simple system consisting of a single spring and a frictional point, then extend our analysis to the Hertzian (parabolic) contact using the Method of Dimensionality Reduction [25]. For simplicity we do not deal with system dynamics, and instead impose a forced oscillation of the indentation depth. This restricts our analysis to systems where the contact stiffness is small compared with the stiffness of the system as a whole and the inertia of the contact region thus does not play any role. An analysis involving system dynamics is published in the second part of this two-part paper.

Another contribution of this paper is the consideration of large oscillation amplitudes, when the indenter starts jumping. To our knowledge this case has not previously been considered in theoretical models.

2 Simplified one-spring model

Let us consider an elastic body that is brought into contact with a flat substrate and then subjected to a superposition of an oscillation in the direction normal to the substrate and movement with a constant velocity in the tangential direction. We will assume that Coulomb's law of friction with a constant coefficient of friction μ_0 is valid in the contact. We first consider a very simple model consisting of a single spring with normal stiffness k_z and tangential stiffness k_x. As the reference state, the unstressed state in the moment of first contact with the substrate is chosen. Let us denote the horizontal and vertical displacements of the upper point of the spring from the reference state by u_x and u_z and the horizontal displacement of the lower (contact) point by $u_{x,c}$. The upper point is forced to move according to

$$u_z = u_{z,0} - \Delta u_z \cos \omega t \quad \text{and} \quad \dot{u}_{\dot{x}} = v_x \tag{1}$$

(see Fig. 1).

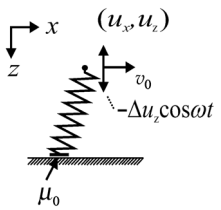


Fig. 1 The simplest model of a tribological contact with a constant contact stiffness represented as a single spring, which has a normal stiffness k_z and a tangential stiffness k_x. The upper end of the spring is forced to move according to Eq. (1). At the lower end (immediate contact spot), Coulomb's law of friction with a constant coefficient of friction μ_0 is assumed.

2.1 Small oscillation amplitudes (no "jumping")

Let us start our consideration with the case of sufficiently small oscillation amplitudes, $\Delta u_z < u_{z,0}$, so that the indenter remains in contact with the substrate at all times. As for the horizontal movement, the lower point of the spring can be either in stick or slip states. During the slip phase the tangential force $f_x = k_x(u_x - u_{x,c})$ is equal to the normal force $f_z = k_z(u_{z,0} - \Delta u_z \cos\omega t)$ multiplied with the coefficient of friction: $k_x(u_x - u_{x,c}) = \mu_0 k_z(u_{z,0} - \Delta u_z \cos\omega t)$. Differentiating this equation with respect to time gives $k_x(v_0 - \dot{u}_{x,c}) = \mu_0 k_z \omega \Delta u_z \sin\omega t$. For the tangential velocity of the lower contact point, it follows that $\dot{u}_{x,c} = v_0 - \mu_0(k_z/k_x)\omega\Delta u_z \sin\omega t$. This equation is only valid when $\dot{u}_{x,c} > 0$, and the foot point of the spring will transition from the sliding state to the sticking state when the condition $\dot{u}_{x,c} = 0$ is fulfilled. This occurs at the time t_1 which satisfies the following equation:

$$\dot{u}_{x,c} = v_0 - \mu_0(k_z/k_x)\omega\Delta u_z \sin\omega t_1 = 0 \tag{2}$$

Introducing a dimensionless velocity

$$\overline{v} = \frac{k_x}{k_z}\frac{v_0}{\mu_0\omega\Delta u_z} \tag{3}$$

we can rewrite Eq. (2) in the form

$$\sin\omega t_1 = \overline{v} \tag{4}$$

For $\overline{v} > 1$, this equation has no solutions, and the spring continues sliding at all times. Since, in this case, the tangential force remains proportional to the product of the normal force and the macroscopic coefficient of friction μ_0 at all times, there is no reduction of the macroscopic force of friction.

For dimensionless velocities smaller than one, $\overline{v} < 1$, Eq. (4) has solutions and the movement of the contact point will consist of a sequence of sliding and sticking phases, where the sliding phase ends at time t_1 given by Eq. (4). The tangential force at this point is equal to $f_x = \mu_0 k_z(u_{z,0} - \Delta u_z \cos\omega t_1)$ or taking Eq. (4) into account:

$$f_x(t_1) = \left(\mu_0 k_z u_{z,0} - \sqrt{\left(\mu_0 k_z \Delta u_{z,0}\right)^2 - \left(\frac{v_x k_x}{\omega}\right)^2}\right) \tag{5}$$

During the sticking stage the tangential force increases linearly according to

$$f_x^{(\mathrm{stick})}(t) = f_x(t_1) + k_x v_x\left(t - t_1\right) \tag{6}$$

The next phase of slip starts at time t_2 when the tangential force becomes equal to the normal force multiplied by the coefficient of friction (see Fig. 2):

$$f_x(t_1) + k_x v_x(t_2 - t_1) = \mu_0 k_z\left(u_{z,0} - \Delta u_z \cos\omega t_2\right) \tag{7}$$

Or taking Eqs. (5) and (4) into account and using the dimensionless variable Eq. (3),

$$\cos\omega t_2 = \overline{v}\left(\omega t_2 - \arcsin\overline{v}\right) + \sqrt{1-\overline{v}^2} \tag{8}$$

The average value of the frictional force during the whole oscillation period can be calculated as follows:

$$\left\langle f_x \right\rangle = \frac{\omega}{2\pi}\left[\int_{t_1}^{t_2} f_x^{(\mathrm{stick})}(t)\mathrm{d}t + \int_{t_2}^{2\pi/\omega + t_1} f_x(t)\mathrm{d}t\right] \tag{9}$$

Divided by the average normal force, this gives the macroscopic coefficient of friction

$$\mu_{\mathrm{macro}} = \left\langle f_x \right\rangle / \left\langle f_z \right\rangle \tag{10}$$

where $\left\langle f_z \right\rangle = k_z u_{z,0}$ in the non-jumping case, which is considered here. The result of numerical evaluation of the macroscopic coefficient of friction, normalized by the local coefficient of friction μ_0 is presented in Fig. 3. It was found that the numerically obtained dependences of the coefficient of friction on dimensionless velocity and amplitude can be approximated very accurately with the following equation:

$$\frac{\mu_{\mathrm{macro}}}{\mu_0} \approx 1 - \frac{\Delta u_z}{u_{z,0}}\left[\frac{3}{4}(\overline{v}-1)^2 + \frac{1}{4}(\overline{v}-1)^4\right] \tag{11}$$

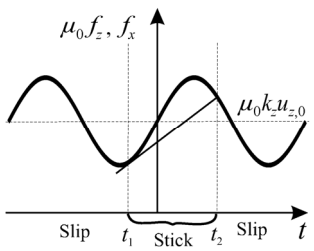


Fig. 2 Schematic presentation of the normal force multiplied with the coefficient of friction (sinusoidal curve) and tangential force (straight line). During the slip phase (before t_1 and after t_2), the tangential force is equal to the normal force times the coefficient of friction, thus both curves coincide. During the stick phase (between t_1 and t_2), the tangential force is smaller than the normal force multiplied by the coefficient of friction. Both forces become equal again at time t_2, where the stick phase ends.

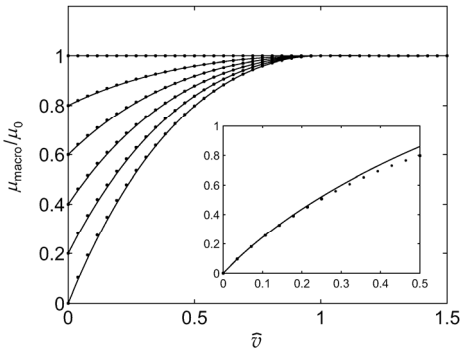


Fig. 3 Dependence of the normalized coefficient of friction on the normalized velocity for $\Delta u_z / u_{z,0} = 0$, 0.2, 0.4, 0.6, 0.8, 1.0 (from top to bottom). Points represent the results of numerical evaluation of the Integral (9). Solid lines represent the empirical Approximation (11). The inset shows the low-velocity asymptotic Solution (12) (solid line) compared to the numerical Solution (9) (points).

A comparison of this approximation with numerical results provided by Eqs. (9) and (10) is shown in Fig. 3. The low-velocity limit of the coefficient of friction can be derived analytically by replacing the time-dependence of the normal force with its Taylor series around the points $\omega t = 0$ and $\omega t = 3\pi/2$ and repeating the above calculations including integration of (9), which provides the result

$$\frac{\mu_{\text{macro}}}{\mu_0} = 1 + \frac{\Delta u_z}{u_{z,0}}\left(-1 + \pi\bar{v} - \frac{4}{3}\sqrt{\pi}\bar{v}^{3/2} + \frac{1}{2}\bar{v}^2\right) \qquad (12)$$

This dependence is asymptotically exact in the limit of small sliding velocities. Like the empirical Approximation (11) it contains only two dimensionless variables: the dimensionless amplitude of oscillation $\Delta u_z / u_{z,0}$ and the dimensionless sliding velocity (3). A comparison with the numerical results is shown for the case of the critical oscillation amplitude, $\Delta u_z / u_{z,0} = 1$, in the inset of Fig. 3.

Equation (11) can be rewritten in a form explicitly giving the average tangential force (force of friction):

$$\langle f_x \rangle = \mu_0 k_z u_{z,0}\left(1 - \frac{\Delta u_z}{u_{z,0}}\left[\frac{3}{4}(\bar{v}-1)^2 + \frac{1}{4}(\bar{v}-1)^4\right]\right) \qquad (13)$$

Note that the *change* of the friction force due to oscillation, as compared with sliding without oscillation, does depend on the amplitude of oscillation, but does *not* depend on the average normal force:

$$\Delta\langle f_x \rangle = -\mu_0 k_z \Delta u_z\left[\frac{3}{4}(\bar{v}-1)^2 + \frac{1}{4}(\bar{v}-1)^4\right] \qquad (14)$$

As will be shown later, this property implies that Eq. (14) is valid for *arbitrarily-shaped* contacts if the oscillation amplitude is small, and k_x and k_z are understood as the incremental tangential and normal stiffness of the contact.

Equation (11) provides a compact representation of the law of friction. However, it is not always convenient for interpretation of experimental results, as the dimensionless velocity (3) is normalized by the amplitude of velocity oscillation and thus the scaling of the velocity depends on the oscillation amplitude. To facilitate the physical interpretation of experimental results it may be more convenient to normalize the velocity using a value that does not depend on the oscillation amplitude. Introducing a new normalized velocity $\hat{v}$ according to the definition

$$\hat{v} = \frac{k_x}{k_z}\frac{v_x}{\mu_0 \omega u_{z,0}} \qquad (15)$$

we can rewrite Eq. (11) in the form

$$\frac{\mu_{\text{macro}}}{\mu_0} \approx 1 - \frac{\Delta u_z}{u_{z,0}}\left[\frac{3}{4}\left(\hat{v}\frac{u_{z,0}}{\Delta u_z} - 1\right)^2 + \frac{1}{4}\left(\hat{v}\frac{u_{z,0}}{\Delta u_z} - 1\right)^4\right] \qquad (16)$$

This dependence is presented in Fig. 4.

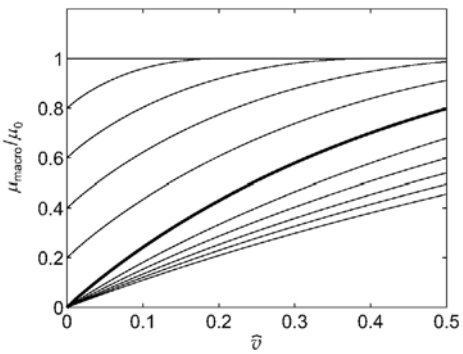


Fig. 4 Dependence of the normalized coefficient of friction on the dimensionless velocity (15): The horizontal line at the constant value 1 corresponds to sliding friction without oscillation. When the oscillation amplitude increases, the static force of friction (at zero velocity) decreases until it vanishes (bold line). This trend is shown in the upper part of the plot for amplitudes $\Delta u_z / u_{z,0} = 0.2$, 0.4, 0.6, 0.8, 1.0 (from top to bottom). Further increase of the oscillation amplitude leads to loss of contact during a part of the oscillation period. In this range of oscillation amplitudes, the static friction force remains zero, and the slope of the dependency decreases with increasing oscillation amplitude. This is shown in the lower part of the plot for amplitudes $\Delta u_z / u_{z,0} =$ 1.2, 1.4, 1.6, 1.8, 2.0 (from top to bottom).

2.2 Large oscillation amplitudes ("jumping")

If the amplitude of normal oscillations exceeds the average indentation depth, the indenter starts to "jump": In this case it will be in contact with the substrate only during part of the oscillation period and in the no-contact-state for the rest of the time. For a jumping contact, analytical considerations become too cumbersome, and we will only present the results of numerical modeling of the behavior of the system. In our model, the movement of the contact point is determined by local stick and slip conditions: As long as the tangential spring force is smaller than the normal force multiplied by the coefficient of friction, the contact point is considered to be stuck to the substrate. If in any particular time step the tangential force exceeds the maximum friction force, it is brought into equilibrium by appropriately changing the contact coordinate. Overall, the system undergoes alternating contact and non-contact phases, while each contact phase may be divided into stick and slip phases. The average force during one complete period of oscillation, divided by the average normal force, results in the macroscopic coefficient of friction, μ_{macro}. It can be shown that, as in the non-jumping case, the dimensionless coefficient of friction μ_{macro}/μ_0 is a function of only the dimensionless velocity $\overline{v}$ given by Eq. (3) and the dimensionless oscillation amplitude. This property was checked by varying (dimensional) system parameters while preserving the values of the two dimensionless parameters. The numerical results for the jumping case are shown in Fig. 5. One can see that the dependence of the reduced coefficient of friction on the reduced velocity does not change significantly after the reduced oscillation amplitude exceeds the critical value 1, where static friction first disappears. Thus, as a very rough approximation, one can use the relation (11) with the critical oscillation amplitude for the whole range of jumping contacts:

$$\frac{\mu_{\text{macro}}}{\mu_0}\approx 1-\left[\frac{3}{4}(\overline{v}-1)^2+\frac{1}{4}(\overline{v}-1)^4\right]\text{(for the jumping case)} \tag{17}$$

It is interesting to note that the critical value of the dimensionless velocity $\overline{v}$, after which there is continuous sliding and the macroscopic coefficient of friction coincides with the microscopic one, is equal to 1 both in the non-jumping and jumping regimes.

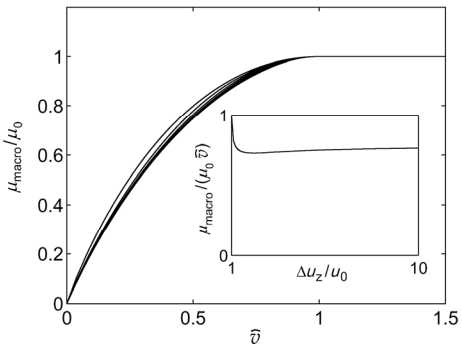


Fig. 5 Dependence of the normalized coefficient of friction on the dimensionless velocity $\overline{v}$ (3) for the "jumping" case, i.e., when the oscillation amplitude exceeds the average indentation depth. Curves are shown for 11 oscillation amplitudes from $\Delta u_z/u_{z,0}=1$ to $\Delta u_z/u_{z,0}=2$ with a step of 0.1. The curves for higher oscillation amplitudes "pile up" towards a limiting curve. The inset shows the dependence of the slope of the low-velocity asymptote (21) on $\Delta u_z/u_{z,0}$. One can see that it depends only weakly on the oscillation amplitude: Once the sample starts jumping the slope drops rapidly by about 20% and then remains practically constant with a limiting value of $\pi/4$.

The low-velocity asymptote of the dependence of the coefficient of friction can be easily found analytically. It is instructive to do this for a better understanding of the details of the dependence and of possible deviations from the rough estimate (17). At sufficiently low velocities, the spring will stick as soon as it comes into contact with the substrate. From Eq. (1), we can see that the times when contact is lost or regained are determined by the equation

$$\omega t_{1,2}=\pm\arccos\left(u_0/\Delta u_z\right),\ \text{for}\ \left|\Delta u_z\right|>u_0 \tag{18}$$

The spring comes into contact in fully relaxed state and is then moved with the constant velocity v_0 during the contact time $t_{\text{contact}}=2\pi/\omega-2t_2$. At low velocities the spring will remain in stick for almost the entire contact time, so that the average tangential force during the contact time can be estimated as $\langle F_x\rangle_{\text{contact}}=k_x v_0 t_{\text{contact}}/2$ and the average tangential force during the whole oscillation period as

$$\langle F_x\rangle=\frac{k_x v_0\omega t_{\text{contact}}^2}{2\cdot 2\pi}=\frac{k_x v_0}{\pi\omega}\left(\pi-\arccos\left(\frac{u_0}{\Delta u_z}\right)\right)^2 \tag{19}$$

The average normal force is given by:

$$\langle F_z \rangle = k_z u_0 \left(1 - \frac{1}{\pi} \arccos\left(\frac{u_0}{\Delta u} \right) + \frac{1}{\pi} \frac{\Delta u}{u_0} \sqrt{1 - \left(\frac{u_0}{\Delta u} \right)^2} \right) \quad (20)$$

with which we finally find the normalized coefficient of friction:

$$\frac{\mu_{\text{macro}}}{\mu_0} = \bar{v} \frac{\left(1 - \frac{1}{\pi} \arccos\left(\frac{u_0}{\Delta u_z} \right) \right)^2}{\frac{u_0}{\Delta u_z} \left(1 - \frac{1}{\pi} \arccos\left(\frac{u_0}{\Delta u_z} \right) + \frac{1}{\pi} \frac{\Delta u}{u_0} \sqrt{1 - \left(\frac{u_0}{\Delta u_z} \right)^2} \right)} \quad (21)$$

This result illustrates once more that the reduced coefficient of friction is a function of only the dimensionless velocity $\bar{v}$ and the dimensionless oscillation amplitude $\Delta u_z / u_{z,0}$. The dependence of the slope of the low-velocity asymptote on the dimensionless oscillation amplitude $\Delta u_z / u_{z,0}$ is shown in the inset of Fig. 5. One can see that when the sample starts jumping the slope drops rapidly by about 20% and then it remains nearly constant at the limiting value of $\pi/4$, thus an explicit expression for the low-velocity asymptote in the jumping regime can be written (in the original dimensional variables) as:

$$\mu_{\text{macro}} \approx \frac{\pi}{4} \frac{k_x}{k_z} \frac{v_0}{\omega \Delta u_z} \quad \text{(low velocity asymptote; } \Delta u_z > u_{z,0}\text{)} \quad (22)$$

As mentioned above, for comparison with experiments it may be preferable to use the dimensionless velocity $\hat{v}$ (15), which does not depend on the oscillation amplitude. In terms of this velocity, the coefficient of friction is shown in Fig. 4 for both jumping and non-jumping regimes, separated by a bold solid line corresponding to the critical amplitude $\Delta u_z / u_{z,0} = 1$. Overall, one can see that an increase of the oscillation amplitude first leads to a decrease of the static coefficient of friction at low sliding velocities. At the critical amplitude, the static coefficient of friction vanishes and remains zero during further increases of the oscillation amplitude, while the overall dependence on velocity starts to "tilt" (the slope of the dependence decreases with increasing oscillation amplitude).

3 Reduction of friction in a Hertzian contact

In Section 2, we considered a simplified model in which it was assumed that the contacting bodies have a constant contact stiffness that does not depend on the indentation depth. This model can be realized experimentally by using a flat-ended cylindrical pin or a curved body with a flat end (e.g., due to wear). However, in the general case the body in contact will have curved or rough surfaces so that the contact stiffness will depend on the indentation depth. In this section we generalize the results obtained in the previous section for more general contact configurations.

In our analysis of the contact of a curved body with the substrate we will use the so-called Method of Dimensionality Reduction (MDR). As shown in Ref. [26], the contact of arbitrarily shaped bodies can be described (in the usual half-space approximation) by replacing it with a contact of an elastic foundation with a properly defined planar shape $g(x)$, as shown in Fig. 6. The elastic foundation consists of a linear arrangement of independent springs with normal stiffness Δk_z and tangential stiffness Δk_x and with sufficiently small spacing Δx. For an exact mapping, the stiffness of the springs has to be chosen according to Refs. [25, 27]:

$$\Delta k_z = E^* \Delta x \quad \text{with} \quad \frac{1}{E^*} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \quad (23)$$

$$\Delta k_x = G^* \Delta x \quad \text{with} \quad \frac{1}{G^*} = \frac{2-\nu_1}{4G_1} + \frac{2-\nu_2}{4G_2} \quad (24)$$

where E_1 and E_2 are the moduli of elasticity, ν_1 and ν_2 the Poisson numbers, and G_1 and G_2 the shear moduli of the bodies. The "equivalent shape" $g(x)$ providing the exact mapping can be determined either

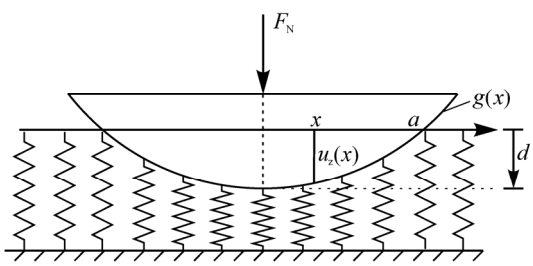


Fig. 6 Schematic presentation of the contact of a transformed planar profile with an effective elastic foundation as prescribed by the rules of the Method of Dimensionality Reduction (MDR).

analytically (e.g., for axis-symmetric profiles [25, 27]), or by asymptotic [26], numerical [26, 28] or experimental methods. It is important to note that an equivalent profile does exist for arbitrary topographies of contacting bodies. Once determined, this equivalent profile can be used to analyze arbitrary dynamic normal and tangential loading histories. If the body is moved tangentially, the same law of friction that is valid for the three-dimensional bodies is applied for each individual spring, using the same coefficient of friction μ_0. If the above rules are observed, the relations between macroscopic properties of the contact (in particular the normal and tangential force-displacement-relationships) will identically coincide with those of the initial three-dimensional problem [26].

3.1 Arbitrary surface topography and small amplitude of oscillations

Let us start by deriving the reduction of friction force for the case of arbitrary contact geometry and *small* oscillation amplitudes. Consider the MDR-representation of the problem in Fig. 6. If the oscillation amplitude is small, then most springs which came into contact with the elastic foundation during the initial indentation by $u_{z,0}$ will remain in contact at all times. Thus, the result (14), which is valid in the non-jumping case, is applicable for most of the springs in the contact; we only have to replace the normal contact stiffness by the stiffness of a single spring:

$$\Delta\left\langle f_x\right\rangle_{\text{one spring}} = -\mu_0 \Delta k_z \Delta u_z \left[\frac{3}{4}\left(\overline{v}-1\right)^2 + \frac{1}{4}\left(\overline{v}-1\right)^4\right] \quad (25)$$

The oscillation amplitude and the expression in the brackets are the same for all springs. Summing over all springs therefore just means replacing the stiffness of one spring by the total stiffness of all springs in contact, k_z, which leads us back to Eq. (14), which is thus generally valid for arbitrary contact shapes.

3.2 Parabolic surface profile and arbitrary amplitude of oscillations

For a parabolic profile $z = r^2/(2R)$ the equivalent plane profil $g(x)$ is given by Ref. [25]: $g(x) = x^2/R$. In our numerical simulations, this profile was first indented by $u_{z,0}$. Subsequently, the indenter was subjected to superimposed normal oscillation and tangential movement with constant velocity according to Eq. (1). Since the springs of the MDR model are independent, the simulation procedure for each spring is exactly as described in Section 2: The movement of the contact point of each spring of the elastic foundation was determined by the stick and slip conditions: as long as the tangential spring force remained smaller than the normal force multiplied by the coefficient of friction, the contact point remained stuck to the substrate. If in a particular time step the tangential force exceeded the critical value, it was reset to the critical value by appropriately changing the contact coordinate. This procedure unambiguously determines the normal and tangential force in each spring of the elastic foundation at each time step. By summing the forces of all springs the total normal and tangential force are determined. After averaging over one period of oscillation, the macroscopic coefficient of friction was found by dividing the mean tangential force by the mean normal force. This coefficient of friction, normalized by the local coefficient of friction μ_0, once again appears to be a function of only two parameters: the dimensionless velocity (either $\overline{v}$ (3), see Fig. 7, or $\hat{v}$ (15), see Fig. 8) and the dimensionless oscillation amplitude $\Delta u_z/u_{z,0}$.

For a parabolic profile, the dependences look qualitatively very similar to those for a single spring (compare with Fig. 3 and Fig. 4). The dependences have two characteristic features: (a) the static force of friction—the starting point of the curve at zero velocity and (b) the critical velocity $\overline{v}=1$ after which there is no further influence of oscillations on the macroscopic coefficient of friction.

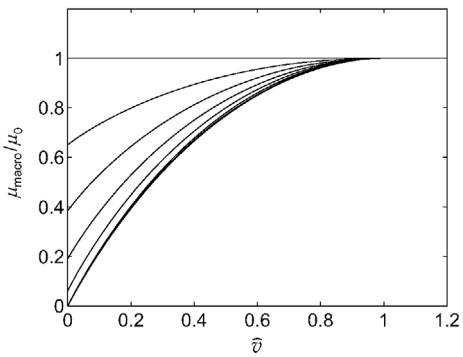


Fig. 7 Dependence of the normalized coefficient of friction on the normalized velocity $\overline{v}$ defined by Eq. (3) for the oscillation amplitudes $\Delta u_z/u_{z,0} = 0,\ 0.2,\ 0.4,\ 0.6,\ 0.8,\ 1.0$ (top to bottom).

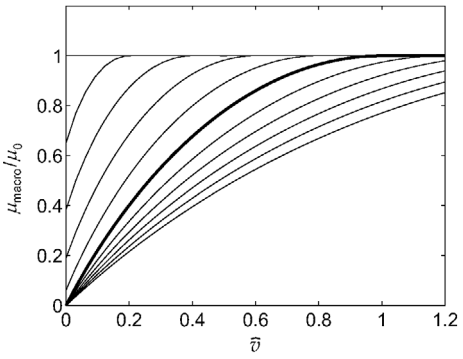


Fig. 8 Dependence of the normalized coefficient of friction on the normalized velocity $\bar{v}$ defined by Eq. (15) for the oscillation amplitudes $\Delta u_z/u_{z,0}=0$, 0.2, 0.4, 0.6, 0.8, 1.0 (the bold line and all curves in the upper-left part) and for $\Delta u_z/u_{z,0}=1.2$, 1.4, 1.6, 1.8, 2.0 (bottom-right part).

4 Discussion

Let us summarize and discuss the main findings of the present study and provide a comparison with experimental results. The structure of the obtained dependences of the macroscopic coefficient of friction on the velocity in the presence of oscillations is simple and contains only two main reference points: the static friction force and the critical velocity. The dependence of the static friction force is extremely simple: it is determined just by the minimum of the normal force during the oscillation cycle. The differences of the static friction force for indenters of different shape will therefore be completely determined by the solution of the corresponding normal contact problem. The second reference point is the critical velocity, which separates the velocity interval where the coefficient of friction does depend on the velocity from the interval where there is no further dependence. This critical point is given by the condition $\bar{v}=1$ or explicitly, in dimensional variables:

$$v_0=\frac{E^*}{G^*}\mu_0\omega\Delta u_z \tag{26}$$

Since Mindlin's ratio E^*/G^* is on the order of unity and $\omega\Delta u_z$ is the amplitude of *velocity oscillation*, this means that the critical velocity is roughly speaking the amplitude of the velocity oscillation multiplied with the coefficient of friction. It is astonishing that this simple result is absolutely universal and is valid for both the non-jumping and jumping regimes and for indenters of arbitrary shape.

Thus, one of the reference points is determined solely by the amplitude of displacement oscillation and the other solely by the amplitude of the velocity oscillation. Between these points, the dependence of the coefficient of friction on sliding velocity is accurately approximated by Eq. (11), which can be rewritten in a universal form that does not depend on the indenter shape:

$$\frac{\mu_{\text{macro}}}{\mu_0}\approx 1-\left(1-\frac{\mu_{\text{static}}}{\mu_0}\right)\left[\frac{3}{4}(\bar{v}-1)^2+\frac{1}{4}(\bar{v}-1)^4\right] \tag{27}$$

The indenter shapes will only influence the static coefficient of friction in the above equation.

For practical applications one can use an even simpler approximation differing from Eq. (27) by 1% or less:

$$\frac{\mu_{\text{macro}}}{\mu_0}\approx 1-\left(1-\frac{\mu_{\text{static}}}{\mu_0}\right)(1-\bar{v})^{2.4} \tag{28}$$

Substituting the definition of $\bar{v}$, we can write this dependence in the initial dimensional variables:

$$\mu_{\text{macro}}\approx\mu_0-(\mu_0-\mu_{\text{static}})\left(1-\frac{G^*}{E^*}\frac{v_0}{\mu_0\omega\Delta u_z}\right)^{2.4} \tag{29}$$

This equation contains in a condensed form all essential results of the present study. Most interestingly, it is approximately valid in both non-jumping and jumping regimes and for all indenter shapes. As long as the amplitude of oscillation is smaller than the average indentation (no jumping), the static friction force decreases monotonously with increasing oscillation amplitude. After reaching the critical oscillation amplitude, the static friction force vanishes and remains zero at larger oscillation amplitudes, but Eq. (27) still remains valid in a good approximation. From the critical amplitude onwards, the force-velocity dependencies start to "tilt".

The described features can be readily recognized in the experimental data shown in Fig. 9.

Comparison of the experimental results with the theoretical predictions in Fig. 4 shows both similarities and differences. For example, we also observe the

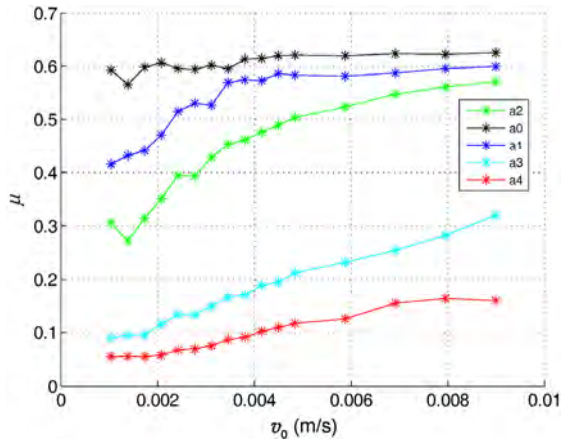


Fig. 9 Experimentally determined dependences of the coefficient of friction on sliding velocity between a steel sphere and a steel disc for increasing amplitudes of out-of-plane oscillation obtained by Milahin (Source: [29], reproduced with permission of the author). The upper-most curve corresponds to sliding in the absence of oscillation. The second, third and fourth curves correspond to amplitudes of $\Delta u_z = 0.06$ μm, $\Delta u_z = 0.10$ μm, $\Delta u_z = 0.16$ μm and $\Delta u_z = 0.27$ μm.

decrease of static friction and subsequent "tilting" of the dependences in the experimental data. Similar behavior was also observed in Ref. [30]. A difference between our theory and experiment is that the static coefficient of friction does not vanish entirely even at large oscillation amplitudes. This effect is known also for other modes of oscillation and is related to the microscopic heterogeneity of the frictional system, which means that Coulomb's law of friction is not applicable at very small space scales [31].

As we noted in the introduction, we considered a case of a soft contact and a rigid measuring system. In the opposite case of a very stiff contact and soft surrounding system, the dependences of the coefficient of friction on the oscillation amplitude appear to be essentially influenced by the inertial properties of the system [16]. An analysis carried out by Teidelt in Ref. [30] has shown that for the measuring system described in Ref. [16] a reasonable agreement between experiment and theory can only be achieved if the contact stiffness is taken into account. However, under other conditions—and in particular depending on the frequency of oscillations—the assumption of soft contact can fail. For such cases, a more general analysis has to be carried out, which is provided in the second part of this series [32].

Acknowledgements

This work was supported in part by the Ministry of Education of the Russian Federation, by COST Action MP1303 and the Deutsche Forschungsgemeinschaft.

References

[1] Eaves A, Smith A, Waterhouse W, Sansome D. Review of the application of ultrasonic vibrations to deforming metals. *Ultrasonics* **13**(4): 162–170 (1975)

[2] Siegert K, Ulmer J. Superimposing ultrasonic waves on the dies in tube and wire drawing. *J Eng Mater Technol* **123**(4): 517 (2001)

[3] Murakawa M. The utility of radially and ultrasonically vibrated dies in the wire drawing process. *J Mater Process Technol* **113**(1–3): 81–86 (2001)

[4] Ashida Y, Aoyama H. Press forming using ultrasonic vibration. *J Mater Process Technol* **187**–**188**: 118–122 (2007)

[5] Fridman H D, Levesque P. Reduction of static friction by sonic vibrations. *J Appl Phys* **30**(10): 1572 (1959)

[6] Pohlman R. Influence of ultrasonic vibration on metallic friction. *Ultrasonics* **4**: 178–185 (1966)

[7] Godfrey D. Vibration reduces metal to metal contact and causes an apparent reduction in friction. *Tribol Trans* **10**(2): 183–192 (1967)

[8] Lenkiewicz W. The sliding friction process—Effect of external vibrations. *Wear* **13**(2): 99–108 (1969)

[9] Tolstoi D M. Significance of the normal degree of freedom and natural normal vibrations in contact friction *Wear* **10**(3): 199–213 (1967)

[10] Schmidt J. A note on the contact problem in an ultrasonic travelling wave motor. *Int J Non-Linear Mech* **31**(6): 915–924 (1996)

[11] Storck H, Littmann W, Wallaschek J, Mracek M. The effect of friction reduction in presence of ultrasonic vibrations and its relevance to travelling wave ultrasonic motors. *Ultrasonics* **40**(1–8): 379–383 (2002)

[12] Socoliuc A, Gnecco E, Maier S, Pfeiffer O, Baratoff A, Bennewitz R, Meyer E. Atomic-scale control of friction by actuation of nanometer-sized contacts. *Science* **313**(5784): 207–210 (2006)

[13] Guo W, Yin J, Qiu H, Guo Y, Wu H, Xue M. Friction of low-dimensional nanomaterial systems. *Friction* **2**(3): 209–225 (2014)

[14] Chowdhury M, Helali M. The effect of amplitude of vibration on the coefficient of friction for different materials. *Tribol Int* **41** (4): 307–314 (2008)

[15] Popov V L, Starcevic J, Filippov A E. Influence of ultrasonic in-plane oscillations on static and sliding friction and intrinsic length scale of dry friction processes. *Tribol Lett* **39**: 25–30 (2010)

[16] Teidelt E, Starcevic J, Popov V L. Influence of ultrasonic oscillation on static and sliding friction. *Tribol Lett* **48**: 51–62 (2012)

[17] Teidelt E, Willert E, Filippov A E, Popov V L. Modeling of the dynamic contact in stick-slip microdrives using the method of reduction of dimensionality. *Phys Mesomech* **15**(5–6): 287–292 (2012)

[18] Nguyen H X, Teidelt E, Popov V L, Fatikow S. Dynamical tangential contact of rough surfaces in stick-slip microdrives: modeling and validation using the method of dimensionality reduction. *Phys Mesomech* **17**(4): 304–310 (2014)

[19] Starcevic J, Filippov A E. Simulation of the influence of ultrasonic in-plane oscillations on dry friction accounting for stick and creep. *Phys Mesomech* **15**(5–6): 330–332 (2012)

[20] Grzemba B, Pohrt R, Teidelt E, Popov V L. Maximum micro-slip in tangential contact of randomly rough self-affine surfaces. *Wear* **309**(1): 256–258 (2014)

[21] Milahin N, Starcevic J. Influence of the normal force and contact geometry on the static force of friction of an oscillating sample. *Phys Mesomech* **17**(3): 228–231 (2014)

[22] Milahin N, Li Q, Starčević J. Influence of the normal force on the sliding friction under ultrasonic oscillations. *Facta Universitatis. Series: Mechanical Engineering* **13**(1): 27–32 (2015)

[23] Popov V L, Dimaki A, Psakhie S, Popov M. On the role of scales in contact mechanics and friction between elastomers and randomly rough self-affine surfaces. *Sci Rep* **5**: 11139 (2015)

[24] Popov V L. What does friction really depend on? Robust governing parameters in contact mechanics and friction. *Phys Mesomech* **19** (2): 115–122 (2016)

[25] Popov V L, Heß M. *Method of Dimensionality Reduction in Contact Mechanics and Friction*. Berlin-Heidelberg (Germany): Springer, 2015.

[26] Argatov I, Heß M, Pohrt R, Popov V L. The extension of the method of dimensionality reduction to non-compact and non-axisymmetric contacts. *ZAMM-J Appl Math Mech* 96(10): 1144–1155 (2016)

[27] Popov V L. *Kontaktmechanik und Reibung. Von der Nanotribologie bis zur Erdbebendynamik*. 3. überarbeitete Auflage. Berlin-Heidelberg (Germany): Springer Vieweg, 2015.

[28] Popov V L, Pohrt R, Heß M. General procedure for solution of contact problems under dynamic normal and tangential loading based on the known solution of normal contact problem. *J Strain Analysis* **51**(4): 247–255 (2016)

[29] Milahin N. Robuste Einflussparameter für Reibung und Oberflächenmodifizierung unter Einfluss von Ultraschall. PhD thesis. Technische Universität Berlin, 2016.

[30] Teidelt E. Oscillating contacts: Friction induced motion and control of friction. PhD thesis. Technische Universität Berlin, 2015.

[31] Dudko O K, Popov V L, Putzar G. Tribospectroscopy of randomly rough surfaces. *Tribol Int* **39**(5): 456–460 (2006)

[32] Mao X, Popov V L, Starcevic J, Popov M. Reduction of friction by normal oscillations. II. In-plane system dynamics. *Friction*, accepted (2017)

Mikhail POPOV. He obtained a bachelor degree in computer science from the Freie Universität Berlin in 2011 and a master degree in engineering science from the Tomsk Polytechnic University and Berlin Technical University in 2016. Currently he is a research assistant at the Tomsk Polytechnic University and a PhD student at the Berlin University of Technology. His research interests include efficient numerical methods in tribology, rubber friction, as well as friction and damping under the influence of oscillations.

Valentin L. POPOV. He is full professor at the Berlin University of Technology, studied physics (1976—1982) and obtained his doctorate in 1985 from the Moscow State Lomonosov University. He worked at the Institute of Strength Physics of the Russian Academy of Sciences. After a guest professorship in the field of theoretical physics at the University of Paderborn (Germany) from 1999 to 2002, he has headed the Department of System Dynamics and the Physics of Friction of the Institute of Mechanics at the Berlin University of Technology. His areas of interest include tribology, nanotribology, tribology at low temperatures, biotribology, the influence of friction through ultrasound, numerical simulation of frictional processes, research regarding earthquakes, as well as topics related to materials science such as the mechanics of elastoplastic media with microstructures, strength of metals and alloys, and shape memory alloys. He has published over 100 papers in leading international journals during the past 5 years. He is the author of the book "*Contact Mechanics and Friction: Physical Principles and Applications*" which appeared in three German, two English, Chinese, and Russian editions and co-author of the book on "*Method of dimensionality Reduction in Contact Mechanics and Friction*" which appeared in German, English and Chinese editions. He is the joint editor of international journals and regularly organizes international conferences and workshops on diverse tribological themes. He is a member of the Scientific Council of the German Tribological Society. He has intensively collaborated with many industrial corporations and possesses experience in implementing the results of scientific research in industrial applications.

Nikita POPOV. Master student in two majors: (a) physics and (b) computer science at the Berlin University of Technology.

Friction **5**(1): 56–65 (2017)
DOI 10.1007/s40544-017-0138-x
ISSN 2223-7690
CN 10-1237/TH
RESEARCH ARTICLE

Correlation of field and experimental test data of wear in heavy commercial vehicle brake liners

B. SURYA RAJAN[1,*], **M. A. SAI BALAJI**[1], **C. VELMURUGAN**[2]

[1] *Department of Mechanical Engineering, BS Abdur Rahman University, Chennai 600048, India*

[2] *Department of Mechanical Engineering, Panimalar Institute of Technology, Chennai 600048, India*

Received: 02 June 2016 / Revised: 06 August 2016 / Accepted: 06 December 2016

Abstract: The correlation between the wear behavior of a heavy commercial vehicle (HCV) brake liner tested under controlled laboratory conditions and that in actual field conditions is investigated. A brake liner study for friction and wear is performed on an inertia brake dynamometer (IBD) at different temperatures (200 ℃, 250 ℃, and 300 ℃) using 6000 brake actuations in a laboratory. The total wear loss of the brake liner at three different temperatures for 6000 brake actuations in IBD is found to be 1.12 mm. The actual field test is conducted on four different HCVs, namely, a city bus (CB), a high speed bus (HSB), a highway truck (HWT), and a tipper lorry (TL). These HCVs run at different terrain/traffic conditions and load conditions. When comparing the predicted life of the brake liner through the IBD test with the actual life of the brake liner in different HCVs, a vast difference is observed. Due to the large variation of liner life observed between the actual and predicted tests, an extensive field test is conducted. In the field test, the liner with identical formulation is fitted in the four types of HCVs. The predicted life of the liner using IBD is then correlated with the field test observation and a correlation factor is determined. Based on this correlation factor, the predicted life of the liner and the achieved life on the HCV are found to be fairly close. This study will be useful to design the friction material formulation and to predict the actual life of the brake liner for various HCVs.

Keywords: heavy commercial vehicle; wear; brake liner; correlation factor; inertia brake dynamometer; friction

1 Introduction

A brake liner designed/formulated for heavy commercial vehicle (HCV) finds application in different types of vehicles, such as passenger vehicles or load carriers. Furthermore, passenger vehicles include city buses (CB) or long route intercity luxury buses (HSB). Moreover, the load carriers can be tipper lorries (TL) used in mines or trucks carrying heavy loads, which usually operate in properly maintained national highways (HWT). Brake liners used in various HCVs will have different life spans if they are used in different terrains, although the formulation is the same. A customer guideline is required to predict the life of the brake liner for a particular type of HCV. Wear life prediction of a brake lining/pad is extremely difficult, unlike other influencing parameters such as friction, fade, noise, and vibration. The wear of a friction lining (brake liner) is an extrinsic property as it depends on the pressure applied on the lining, temperature increase, speed of the drum, frequency of usage of the brakes by the driver, maintenance of the brake system, and road conditions. Wear formation in a drum brake system is shown in Fig. 1. Whenever the brake is applied, wear particles flow along the direction of rotation of the drum and also partially out of the drum. The flow of wear particles depends on the increase in temperature [1] at the drum and liner interface. The temperature increase depends on numerous parameters such as deceleration, speed,

* Corresponding author: B. SURYA RAJAN, E-mail: suryarajan.b@gmail.com

List of symbols

WL_{Total}	Total wear loss of brake liner thickness in mm after 6000 brake applications	$D_{Act\ 6000\ Brake\ appl.(CB)}$	Actual distance travelled by the city bus in km for 6000 brake application
WL_{Avg}	Average wear loss of brake liner thickness in mm for 2000 brake applications	$D_{Act\ 6000\ Brake\ appl.(HSB)}$	Actual distance travelled by the high speed bus in km for 6000 brake application
WL_{200}	Wear loss of brake liner thickness for 2000 brake application at the rotor temperature of 200 ˚C	$D_{Act\ 6000\ Brake\ appl.(HWT)}$	Actual distance travelled by the highway truck in km for 6000 brake application
WL_{250}	Wear loss of brake liner thickness for 2000 brake application at the rotor temperature of 250 ˚C	L_{IBD}	Predicted life of the brake liner in km through inertia brake dynamometers
WL_{300}	Wear loss of brake liner thickness in mm for 2000 brake application at the rotor temperature of 300 ˚C	L_{FO}	Calculated life of the brake liner in km through field observation
t_{usable}	Usable thickness of the brake liner in mm	L_{Act}	Actual life of the brake liner fitted in the HCV
$D_{6000\ Brake\ appl.(IBD)}$	Distance travelled in km for 6000 brake application considering laboratory standards.	$CF_{(A-B)}$	Correlation factor based on L_{FO} and L_{IBD}
$D_{Act\ 6000\ Brake\ appl.(TL)}$	Actual distance travelled by the tipper lorry in km for 6000 brake application	$CF_{(A-C)}$	Correlation factor based on L_{Act} and L_{IBD}

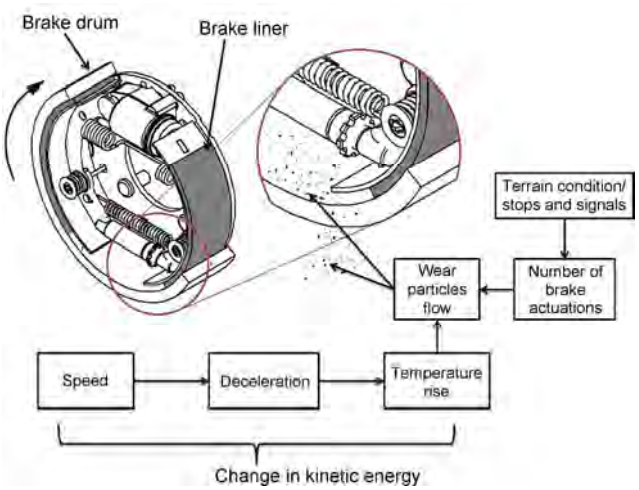


Fig. 1 Schematic diagram of wear formation in a drum brake.

pressure, and terrain/traffic conditions. A wear test is conducted in a typical schedule using an inertia brake dynamometer (IBD) by the original equipment manufacturers (OEMs) in laboratory conditions. The test results of wear conducted in IBD provide guidelines to OEMs to assess the life of the brake liner. Mostly, the same liner formulation is used in numerous HCVs, such as CB, HSB, HWT, and TL. Usually, the predicted life using IBD does not match with the actual life of the brake liners in various HCVs [2].

Research on wear modeling and the prediction of wear life has been performed previously [3]. Therefore, to date, numerous wear models have been proposed for conflicting situations. A linear wear model was developed by Archard [4]. In this model, the wear in terms of volume per sliding distance is found in terms of wear coefficients, which can be influenced by the real contact area, material hardness, pressure distribution, and temperature increase. Alternatively, a nonlinear wear model for the brake pad assembly was developed by Rhee [5] and his influential work was followed by a few other researchers [6, 7]. The estimation of wear of the brake blocks in a short train route with two different modeling approaches using the wear rates from the experimental pin on-disc study is discussed in Ref. [8].

A grey model was developed using the Levenberg-Marquardt (LM) algorithm to predict the life of the brake pads to improve the accuracy of the life prediction

[9]. Estimation of the life of the brake lining on the basis of controlled performance testing of analogue units was explained and the effect of the disc air duct width was calculated to obtain the brake lining service life [2, 10]. A vehicle road test was conducted to obtain the initial braking velocity, temperature, deceleration, and brake pad wear. Based on the road (field) test, a mathematical model was developed and used for modifying the IBD test procedure to check the brake judder and noise. The test results conducted on a modified IBD test schedule coincided fairly well with the road vehicle test [11].

A micro-computer based data logging system was developed in order to study the effect of brake usage on various road surfaces under diverse conditions [12]. A mathematical model was developed to analyze the vehicle braking process with different types of road surfaces and with various vehicle parameters [13]. The formulae that were developed by most of the researchers cited above are more applicable to brake devices in railways and hoists rather than to automotive brake mechanisms, in particular to HCVs. Moreover, most of the models developed by the researchers are based on the results from the experimental studies in a controlled environment. However, the root cause of the present study is due to large variations in brake liner wear life, observed between the experimental (IBD) and field trial (actual) of a typical formulation.

In this investigation, the study of the frictional performance and wear life of a non-asbestos brake liner was conducted using IBD and compared with the actual life of a brake liner fitted in various HCVs operating over different terrain/traffic conditions and load conditions. Whenever a new brake liner formulation requires study of the wear life, the OEM cannot test it on various HCVs. However, this study can provide a few guidelines to predict the life of the brake liner for each HCV based on the IBD wear test results. Furthermore, the current work was performed to understand the wear behavior pattern of non-asbestos friction materials in the laboratory and to correlate the same data with actual field conditions.

2 Experimental setup

A brake liner is formulated using various types of asbestos-free fibers by a combination of Kevlar, steel wool, and glass fiber. This combination is studied for the friction performance and wear properties using the IBD shown in Fig. 2, which has a maximum braking torque capacity of 19,620 N·m. The IBD test is conducted using a computerized control system as per the tailor-made test schedule. A brake liner made by a typical formulation is chosen, and the friction performance, fade/recovery, and wear properties are studied. Performance parameters such as the coefficient of friction, the braking torque generated, and the fade are not considered for the correlation study. The results pertaining to the brake liner wear are considered for this life prediction study.

2.1 IBD

Brake liners are tested on a double-ended IBD, as shown in Fig. 2. This IBD has a 300 hp DC drive capable of accelerating the brake drum to a speed up to 130 km/h. The inertia load can be varied from 60 to 320 N. The test is conducted with the brake assembly and a computer-controlled air brake system capable of generating a maximum pressure of 1.0 bar to 12 bar. Parameters controlled during the test include speed, deceleration, pressure, torque, brake temperature, cooling air speed, and air temperature.

Brake liners are tested using a tailor-made test schedule to simulate the actual condition of wear test procedure, which is similar to JASO C427. This tailor-made wear test procedure recommends testing the brake liners repeatedly at three temperatures, nominally 200 °C, 250 °C, and 300 °C. The temperature of the drum surface is measured by a contact type temperature sensor located exactly in the middle of the inner rotating surface and the liner temperature is measured by the thermocouple placed on the brake liner 3 mm

Fig. 2 Inertia brake dynamometer (IBD).

from the drum contact surface in the radial direction [14]. After each brake actuation, the temperature is maintained at a specific value. In IBD, braking is done against an inertial mass that represents the vehicle inertia. The brake temperature is controlled and adjusted by passing air over the brake drum and liner interface at the interval of braking events using air blowers. In some cases, it is necessary to increase the initial wheel speed from 50 to 90 km/h to obtain the 300 ℃ test temperature.

2.2 Friction performance of the brake liner on the dynamometer test

IBD has inertia masses (wheels) in its section that are rotated by the 300 hp main drive. These inertia masses are used to simulate the kinetic energy of the running or moving HCVs. Depending upon the inertia requirement, the wheels may engage or disengage with the main drive shaft. The function of the main drive is to rotate the inertia masses to a specific speed mentioned in the schedule. After driving the inertia masses at a specified speed, the brake system starts the process of brake actuation to reduce the speed of the inertia masses to zero. The friction performance test is conducted between 35 km/h and 110 km/h. The braking torque (N·m) values are acquired at the different speeds by applying various pressures (between 1 bar and 8 bar) in the brake actuation, as shown in Fig. 3. The torque on the drive axle is monitored for different pressures and the coefficient of friction is obtained.

2.2.1 Discussion of friction behavior of brake liner

Although wear prediction of the brake liner is the main theme of this study, a friction performance test of the chosen brake liner formulation is initially conducted. The graph in Fig. 4 shows that the coefficient of friction continued to be relatively constant at a value of 0.4 until the pressure reached 5 bar and 12,000 N·m torque. As the pressure increases above 5 bar, the coefficient of friction decreases dramatically. This test showed that the lining materials are sensitive to pressure above 5 bar. Moreover, it is observed that for 8 bar pressure at various speeds from 35 km/h to 110 km/h, the coefficient of friction, μ, decreases with increasing speed. This reveals the inconsistency of μ

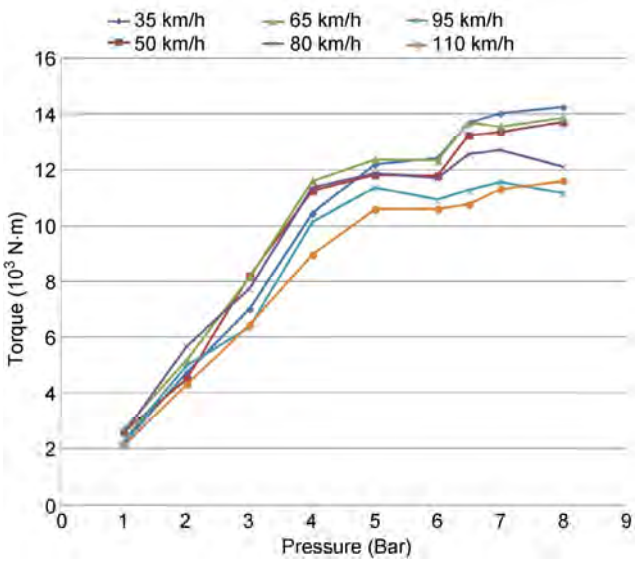


Fig. 3 Braking torque vs pressure (1–8 Bar) at different speeds of the inertia wheel (35–110 km/h).

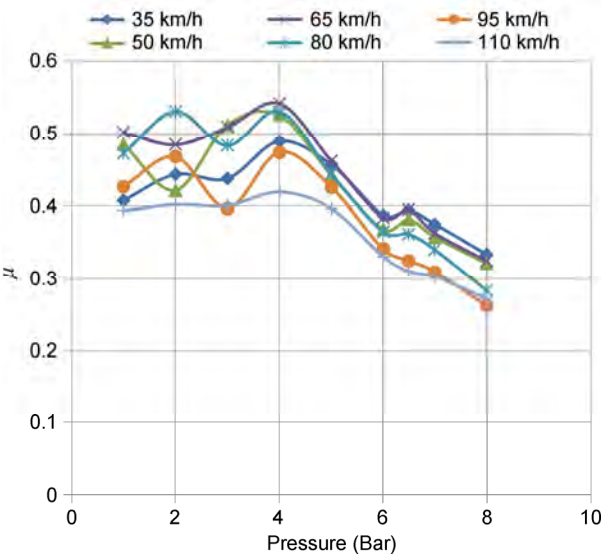


Fig. 4 Coefficient of friction (COF) vs pressure (1–8 Bar) at different speeds of the inertia wheel (35–110 km/h).

with respect to speed. This combined effect of pressure and speed causes the temperature rise at the interface, which decreases the heat resistance components in the formulation. This temperature rise causes the destruction and reformation of the tribo layer, which is responsible for maintaining the stable coefficient of friction [1].

Hence, as observed from Fig. 4, the coefficient of friction developed by the liner is relatively consistent for different pressures and hence acceptable as per

industrial norms. As the performance parameters correlate well with the OEM standards, it is desirable to check another important property, namely, wear as the customer wants their product to have a longer life.

2.3 Wear test of brake liner on IBD

After conducting the friction performance test, the test for wear is conducted using the IBD with 6000 brake actuations. The wear test is conducted for 2000 brake applications by maintaining the brake drum at three different temperatures, namely, 200 ˚C, 250 ˚C, and 300 ˚C. As per JASO C 407 (a wear test schedule for brake liners widely used in Asian countries), the wear test is conducted at 100 ˚C, 200 ˚C, 250 ˚C, and 300 ˚C [14] (350 ˚C is optional, if adequate wear is not obtained). In all four types of HCVs, the interface temperature is never close to 100 ˚C so this temperature is not considered in the tailor-made schedule for correlation purposes. The initial thickness of the liner is measured at 16 different points along the leading side as well as the trailing side of the brake liner. After measuring the thickness at different points on the brake liner, the wear test is conducted for 2000 brake applications at 200 ˚C. The thickness is measured on the same point positions after 2000 brake applications for calculating the thickness loss (WL_{200}). The brake liner wear loss at the end of the test is determined by considering the average wear of both the leading and trailing side to be 0.257 mm (WL_{200} = 0.257 mm) and the corresponding drum wear is 0.019 mm. Moreover, the same sequence of operation is repeated to measure the average wear loss of the brake liner at the rotor temperature of 250 ˚C (WL_{250}) and 300 ˚C (WL_{300}), and the results are shown in Fig. 5. The average wear loss of the liner and drum at 250 ˚C is 0.372 mm (WL_{250} = 0.372 mm) and 0.024 mm, respectively. Similarly, the average wear loss of the brake liner and brake drum at 300 ˚C is 0.491 mm (WL_{300} = 0.491 mm) and 0.029 mm, respectively. The wear characteristics of the liner alone are shown in Fig. 5.

$$WL_{TOTAL} = WL_{200} + WL_{250} + WL_{300} \quad (1)$$

Hence, the total wear loss, WL_{TOTAL}, of the liner after 6000 brake applications is 1.12 mm. The average wear loss in mm considering the different temperatures (200 ˚C, 250 ˚C, and 300 ˚C) for 2000 brake applications

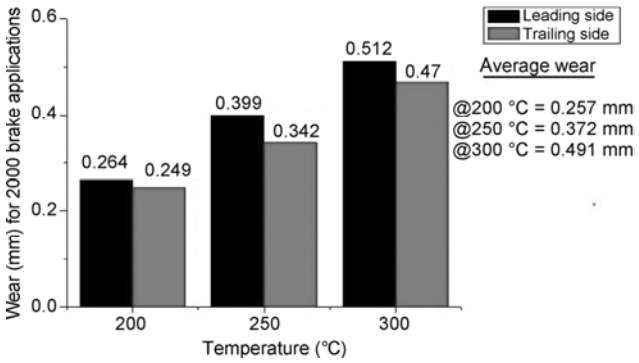


Fig. 5 Wear loss of brake liner in mm at 200 ˚C, 250 ˚C, and 300 ˚C measured in both leading and trailing side.

has been determined as 0.373 mm, calculated using the following equation:

$$WL_{AVG} = \frac{(WL_{200} + WL_{250} + WL_{300})}{3} \quad (2)$$

Based on industrial practices, it is understood that the vehicle will cover a minimum of 1.5 km for each brake application, irrespective of the speed of the vehicle. Therefore, the IBD for 6000 brake applications on the vehicle could cover a distance of 9,000 km. Hence, it is accepted that the total wear loss of the liner is 1.12 mm for 9,000 km of travel for the vehicle. To calculate the total predicted life of the brake liner, L_{IBD} in km traveled, we use

$$L_{IBD} = \frac{(t_{usable}) \times (D_{6000\ Brake\ appl.\ (IBD)})}{WL_{TOTAL}} \quad (3)$$

L_{IBD} is the expected life of the brake liner based on the IBD test result using the industry practice standards. The usable liner thickness (t_{usable}) of the brake liner is 12.1 mm. The total distance traveled for 6000 brake applications, $D_{6000\ Brake\ appl.\ (IBD)}$, is 9,000 km. Hence, the expected life of the brake liner through IBD, L_{IBD}, is 97,232 km. However, in a real scenario, the life of liners used in HWT is more than the expected life predicted through the IBD. In the cases of CB, HSB, and HWT, the actual life (L_{Act}) is very less compared to the predicted life, as shown in Table 1.

It is logical that different vehicles and road conditions cause different working modes for brakes toward the pressure applied. Some other interesting characteristics of the brake loading conditions based on field studies are the work done by the brake, number of brake

Table 1 Comparison of liner life expected through IBD wear test & the actual life of the liner L_{Act}.

Test carried on vehicles	Dynamometer (km)	City bus application (km)	High speed bus application (km)	High way trucks application (km)	Tipper application (km)
Observed liner life in km	97,232	28,000–32,000	15,000–20,000	185,000–235,000	10,000–12,000
Expected through IBD test, L_{IBD}		----------97,232 ----------			
Average actual life, L_{Act}		30,000	17,500	210,000	11,000

applications per km, deceleration, and average temperature at the interface. A detailed survey is conducted to find a number of brake applications in various transport vehicles and based on the survey, in coordination with the drivers, the brake actuations per 100 kilometers are observed and listed in Table 2. The average temperature is measured using an infrared thermometer (Nicety ST530$^+$), which is duly calibrated.

3 Wear test on field

The various categories of HCV for which the field tests were conducted include the following:

(1) City bus (CB)—travels on city roads characterized by very high traffic with a large number of traffic signals and stops.

(2) High speed bus (HSB)—travels on national highways and between cities.

(3) Highway trucks (HWT)—travel on highways with an average speed less than 40 kmph.

(4) Tipper lorry (TL) (Mining operations)—travels in mountain areas, roads characterized by high slopes, sharp curves, high hazardous environments (heat), and roads largely covered with mud.

In India, the road conditions have large variations. As per the OEMs suggestion, the joint field trial is always recommended to obtain the average life of the friction material formulation.

3.1 Observations during field trial and discussion

The wear rate comparison of the brake liners on different HCV applications and different locations are given in the subsequent sections.

3.1.1 CB application

The life of the brake liner in a CB varies from 28,000 to 32,000 km and 30,000 km on average (L_{CB} = 30,000 km), as mentioned in Table 1 [data obtained from Chennai and Cochin]. In the case of the CB, the frequency of brake usage (280 brake applications in 100 km traveled, ref. Table 2) is more due to heavy traffic and number of bus stops and signals. However, the speed of these buses rarely exceeds 50 km/h, with the average speed of 35–40 km/h. The gross weight (vehicle + 70 passengers) of the CB along with the maximum passengers is around 20.80 tons. The temperature at the interface is monitored using an infrared thermometer at regular intervals. The temperature seldom exceeds 220 °C. On average, the deceleration rate is less than 0.3 g. A higher number of brake usages accounted for the wear rate of CB brake liners. The life of a brake liner in Chennai is around 32,000 km, whereas in Cochin, it is around 28,000 km due to a lot of sloppy roads.

Table 2 Observations during the field trials.

Characteristics	City bus	High speed bus	Highway trucks	Tipper mining
Work done by brake (kJ/(t·km))	56.4	39.7	17.6	94.3
Number of brake actuation for 100 km distance travel	280	190	42	390
Number of brake actuation per km	2.8	1.9	0.42	3.9
Actual Distance travel for 6000 brake application, $D_{Act\ 6000\ Brake\ appl.}$ (km)	2143	3158	14286	1552
Deceleration rate in terms of g (average)	0.2	0.4	0.2	0.5
Temperature at interface (°C) (average of the observed values)	193	276	150	285

3.1.2 HSB application

The life of a brake liner in a HSB varies from 15,000 to 20,000 km and 17,500 km on average ($L_{\text{Act(HSB)}}$ = 17,500 km). These data were obtained from the vehicles plying between different towns in the state of Tamil Nadu and Kerala. With regard to the bus operations, as observed in the busy route between two towns, the stopping strategy employed by the drivers was found to be a major factor. The driver's attitude influences the brake temperature and controls the brake liner wear. For the field test, it is observed that the temperature increases significantly when the driver applies the brake suddenly with less stopping distance. The rate of deceleration is high (around 0.5 g) and it has great influence on the temperature increase at the interface, as indicated in Fig. 1. If the driver slows down gradually with a longer stopping distance, the temperature registered is less compared to previous approaches; similar observations are reported elsewhere [15]. Moreover, the driver has to apply some panic braking due to the unexpected driving behavior of the two wheeled vehicles. A higher deceleration rate is the reason for greater wear in the case of HSB when compared to that of the CB. The number of brake applications observed in HSB for 100 km is 190. The gross weight (vehicle + 45 passengers) of the HSB along with the maximum passengers is around 18.90 tons.

3.1.3 HWT application

The life of the brake liner in a HT varies from 185,000 to 235,000 km and 210,000 km on average ($L_{\text{Act (HWT)}}$ = 210,000 km) [data obtained from the vehicles plying between Chennai and Mumbai, which is around 1,700 km]. These HWTs normally run on express highways with minimum traffic interference. Maximum speed reached by these vehicles is around 40 to 50 km/h because of heavy loading (up to 20 tons). They seldom use their brakes. The average number of braking applications for 100 km distance travel in HWT is 42. Hence, the life of the brake liner used in HWT can go more than 200,000 km.

3.1.4 TL application in mining

The life of the brake liners in TLs varies from 10,000 to 12,000 km ($L_{\text{Act(TL)}}$ = 11,000 km) [data obtained from the vehicles at Goa, Hubli, and Mangalore mines]. The average temperature at the interface is 250 ˚C due to the environment, under which these vehicles are operating. The mines usually have higher temperatures than the ambient temperature. The terrain conditions are very rough, hazardous and entirely different compared to the road conditions. The number of brake actuations is higher than other vehicles like CB, HSB, and HWT. Usually, these trucks are used to carrying double the load of buses. These factors contribute to the accelerated wear of the brake liners.

4 Life of brake liner based on inputs from the field observation (L_{FO})

In the IBD schedule, the number of brake applications, decelerations, and temperature increases follow a standard pattern. But based on the observations, the number of brake applications, decelerations, and temperature increases vary for different vehicles and terrain conditions. The actual distance traveled by the HCV ($D_{\text{Act 6000 Brake appl. (HCV)}}$) is found and tabulated in Table 2.

4.1 Tipper lorry (TL)

First, the case of the TL is selected as it shows a large variation to the order of the second degree. Authors substituted the observed number of brake actuations in the field study, i.e., 3.9/km (from Table 2) and the wear rate that occurred at 300 ˚C, i.e., 0.491 mm in IBD, in order to calculate the life of the brake liner based on the field observation ($L_{\text{FO (TL)}}$) for the TL. The average temperature observed at the interface in the TL is about 285 ˚C, which is nearer to 300 ˚C. Hence:

$$L_{\text{FO (TL)}} = \frac{t_{\text{usable}} \times D_{\text{Act 6000 Brake appl. (TL)}}}{3\text{WL}_{300}} \tag{4}$$

The life of the brake liner, based on the field observation $L_{\text{FO (TL)}}$, is 12,749 km, which is very much closer to the actual life of the brake liner fitted in the TL. Similarly, the life of the brake liners based on the field observations for other types of HCVs is listed below.

4.2 City bus (CB)

The CB applied 280 brakes over 100 km (refer to

Table 2) and the temperature increase measured in the interface is less than 200 ℃. Thus,

$$L_{\mathrm{FO\,(CB)}} = \frac{t_{\mathrm{usable}} \times D_{\mathrm{Act\ 6000\ \ Brake\ appl.\ (CB)}}}{3\mathrm{WL}_{200}} \qquad (5)$$
$$L_{\mathrm{FO\,(CB)}} = 33,632 \text{ km}$$

This value is close to the actual life from the field trial, namely 30,000 km.

4.3 High speed bus (HSB)

The HSB applied 190 brakes over 100 km (refer Table 2) and the temperature at the interface is around 300 ℃ due to very high deceleration involved in this route.

$$L_{\mathrm{FO\,(HSB)}} = \frac{t_{\mathrm{usable}} \times D_{\mathrm{Act\ 6000\ \ Brake\ appl.\ (HSB)}}}{3\mathrm{WL}_{300}} \qquad (6)$$

The life of the brake liner based on the field observation $L_{\mathrm{FO\ (HSB)}}$ is 25,942 km, which far exceeds the actual life from the field trial, namely, 17,500 km. The reason for this huge difference needs a detailed investigation.

4.4 Highway truck (HWT)

The HWT applies 42 brakes over 100 km (refer to Table 2) and the temperature increase at the interface is less than 200 ℃. Thus,

$$L_{\mathrm{FO\,(HWT)}} = \frac{t_{\mathrm{usable}} \times D_{\mathrm{Act\ 6000\ \ Brake\ appl.\ (HWT)}}}{3\mathrm{WL}_{200}} \qquad (7)$$

The life of the brake liner based on the field observation $L_{\mathrm{FO(HWT)}}$ is 224,203 km, which is close to the actual life from the field trial, namely 210,000 km. (refer to Table 1).

5 Correlation factor to correlate wear properties on dynamometer and actual life

As observed from Table 3, a similarity exists between the actual life (L_{Act}) of the brake liner and the expected life of the brake liner based on the inputs from the field observation (L_{FO}). Hence, those inputs are provided in Table 4 for modifying the test schedule. The field inputs are always necessary in considering the difference in the standard schedule (in the laboratory test) and the real operating parameters that are observed for different HCVs. For example, in the case of the TL segment, the number of brake applications is around 5.85 per 1.5 km where as in the IBD schedule it is assumed as only one brake application per 1.5 km. Similarly, the normal average temperature at the interface is around 285 ℃, whereas in the IBD schedule it is split up into three different temperature regimes.

Table 3 Correlation factor based on actual life (L_{Act}) and life based on field observation (L_{FO}) vs predicted life through IBD (L_{IBD}).

HCV	(A) L_{IBD} (km)	(B) L_{FO} (km)	(C) L_{Act} (km)	Correlation factor (CF)	
				$\mathrm{CF_{(A\text{-}B)}}$	$\mathrm{CF_{(A\text{-}C)}}$
CB	97,232	33,632	30,000	0.345	0.308
HSB		25,942	17,500	0.266	0.180
HWT		224,203	210,000	2.305	2.150
TL		12,749	11,000	0.131	0.113

(A) L_{IBD} = Expected life of liner based on dynamometer test as per industrial practice (km)
(B) L_{FO} = Expected life of liner based on the inputs from field observation (km)
(C) L_{Act} = Actual life of the liner (km)

Table 4 Inputs to OEM based on field trial.

HCV/IBD	Number of brake actuations	Deceleration	Average speed	Avg. number of brake actuations for 1.5 kms	Deceleration ($\mathrm{m/s^2}$)	Field test temperature (℃)
CB	High	Low	Low	4.2	0.2 g	193
HSB	Low	High	High	2.85	0.4 g	276
HWT	Low	Low	Low	0.62	0.2 g	150
TL	High	High	Low-medium	5.85	0.5 g	285
IBD	High-low	High-low	High-low	1	0.3 g	200, 250 & 300

The correlation factors (A-B) and (A-C) are given by

$$CF_{(A-B)} = \frac{L_{FO(HCV)}}{L_{IBD}} \quad (8)$$

$$CF_{(A-C)} = \frac{L_{Act(HCV)}}{L_{IBD}} \quad (9)$$

Using Eqs. (8) and (9), the correlation factor for wear life is calculated and tabulated in Table 3. To check the dependability of the correlation factor, the brake liners were retested in the CB and found that the brake liner covered 30,420 km (using the calculation, it is 33,632 km). Similarly, the brake liners were tested in the TL and found that the brake liner covered 10,200 km (using the calculation, it is 12,749 km). The predicted wear coincided fairly closely with the measured wear. With this correlation factor,

(1) It is understood that instead of using the same formulated brake liners for different HCV applications, distinct formulations can be obtained for category of vehicles (considering the huge volume of vehicles in each category, it is quite economical).

(2) Else, it is preferable to go for the different test schedule for each vehicle category.

(3) The life of the brake liner can be predicted for the customer when fit into a particular type of HCV.

It is observed that the usage of the brakes by the drivers, tested in the selected routes, has no correlation or resemblance with the standard test schedule followed in the industry. Yet, the formulators expect the wear rate of brake liners to coincide in real scenarios in the same way it performed in the controlled environmental conditions. Hence, some inputs for modification of the IBD schedule are provided. Even then, a larger number of varying parameters make it difficult to accurately predict the wear rate of automobile brake systems.

6 Conclusion

- The performance result tested on the IBD matches the respective vehicle applications. No adverse details were reported by the drivers.
- The life-span of the brake liners tested on the IBD and the field is varied. But, it can be correlated based on deceleration, maximum temperature seen and the number of stops per 100 km of running.
- There is no direct correlation between the wear rate tested on the IBD and the actual field conditions, but mathematical guidelines can be achieved. The IBD test is carried out to reconfirm with the same formulation, and found the wear loss reported on IBD is 1.20 mm (previously 1.12 mm).
- The brake liners tested in the CB covered 30,420 km (using the calculation, it is 33,632 km). Similarly, the brake liners were tested in the TL application and found that the brake liner covered 10,200 km (using the calculation, it is 12,749 km). The predicted wear coincided fairly closely with the measured wear, greater than 80%.
- Deceleration and the corresponding temperature increase mainly contributed to wear as all the brake liners were made by the same formulation, produced by the same manufacturer.
- The deceleration rates, maximum temperatures, and number of brake applications employed by the driver under various conditions should be documented in detail, as these affect the brake liner's life to change them even to a high order of magnitude; and can thus be used in the design stage. There is an opportunity to design materials for each application separately, to provide a higher life-span for each scenario. It is possible to improve the correlation between the test vehicle data and the IBD data; it is extremely unlikely that IBD testing will ever fully replace vehicle testing because of factors such as road conditions, driver usage of brakes, maintenance of brake systems, and auto slack adjuster functioning.

Acknowledgement

The authors are also grateful to Mr. Sanmugam, Asst. Engineer, PRTC corporation, Puducherry for providing the necessary facilities to take field trial.

References

[1] Ostermeyer G P. Friction and wear of brake systems. *Forschung im Ingenieurwesen* **66**(6): 267–272 (2001)

[2] Revin A A, Tyurin S V, Fedotov V N. The way to estimate the lifetime of transport vehicle brake pad according to the results of prototype test exploitation. *Transport* **2:** 27–29 (2007)

[3] Revin A A, Tyurin S V, Fedotov V N, The way to estimate the durability of brake pads according to the results of prototype test exploitation. *Transport* **8:** 38–39 (2007)

[4] Archard J F. Contact and rubbing of flat surfaces. *J Appl Phys* **24**(8): 18–28 (1953)

[5] Rhee S K. Wear mechanisms for asbestos-reinforced automotive friction materials. *Wear* **29**(3): 391–393 (1974)

[6] Pavelescu D, Musat M. Some relations for determining the wear of composite brake materials. *Wear* **27**: 91–97 (1974)

[7] Hohmann C, Schiffner K, Brecht J. Pad wear simulation model. *SAE Technical Paper* No: 1999-01-3392 (1999)

[8] Vernersson T, Lunden R. Wear of brake blocks for in-service conditions-Influence of the level of modelling. *Wear* **314**: 125–131 (2014)

[9] Liu F, Liang X, Ren Q-Z, Zheng J. Life prediction of EMU brake pad based on optimized neural network and grey theory. *International Journal of Digital Content Technology and its Applications* **7**(10): 118 (2013)

[10] Blau P J. Fifty years of research on the wear of metals. *Tribol Int* **30**(5): 321–331 (1997)

[11] Wang Q, Qi G, Zhang G, Pu X. Study on brake durability dynamometer experimental method for brake NVH and wear. *SAE Technical Paper* No. 2014-01-2520 (2014)

[12] Minegishi H, Shimizu H, Wakamatsu H, and Yoshino Y. Prediction of brake pad wear/life by means of brake severity factor as measured on a data logging system. *SAE Technical Paper* No. 840358 (1984)

[13] Bogdevicius M, Vladdimirov O. Efficiency of a braking process evaluating the roughness of road surface. *Transport* **21**(1): 3–7 (2006)

[14] *Wear Test Procedure on Inertia Dynamomete─Brake Friction Materials─JASO C* 427–83. Society of Automotive Engineers of Japan, 1983.

[15] Videla D, Fancher P. Prediction of brake temperatures on urban bus routes. UMTRI 90-32, Engineering Research Division (1990)

B. SURYA RAJAN. He received his bachelor degree in mechanical engineering in 2010 and master degree in engineering design in 2014 from College of Engineering Guindy, Anna University, Chennai, India. Currently he is a Ph.D. student in the B.S. Abdur Rahman University, Vandalur, Chennai, India. His research area includes tribology behavior study of friction composites.

M. A. SAI BALAJI. He is an associate professor in the Department of Mechanical Engineering of BS Abdur Rahman University, Vandalur, Chennai, India. He obtained his bachelor degree in production technology from Anna University, master degree in computer aided design from Madras University and Ph.D. degree in the topic of "Effect of organic components in disc brake pad formulation in relation to fade and wear". His current research area is friction composites.

C. VELMURUGAN. He received his bachelor degree in mechanical engineering and master degree in engineering design from Anna University, Chennai, India. Currently he is working as assistant professor in the Department of Mechanical Engineering, Panimalar Institute of Technology, Chennai, India.

Friction **5**(1): 66–86 (2017)
DOI 10.1007/s40544-017-0139-9
ISSN 2223-7690
CN 10-1237/TH
RESEARCH ARTICLE

Optimization of friction and wear characteristics of varied cryogenically treated hot die steel grade AISI-H13 under dry condition

Sanjeev KATOCH[1,2,*], **Rakesh SEHGAL**[3], **Vishal SINGH**[1]

[1] *Centre for Materials Science and Engineering, National Institute of Technology, Hamirpur (HP) 177005, India*

[2] *Institute for Auto Parts & Hand Tools Technology, A-9, Phase V, Focal Point, Ludhiana (Punjab) 141010, India*

[3] *Department of Mechanical Engineering, National Institute of Technology, Hamirpur (HP) 177005, India*

Received: 30 May 2016 / Revised: 07 September 2016 / Accepted: 11 December 2016

Abstract: Cryogenic treatment (CT) is a relatively new field, which has emerged during the last three decades of the twentieth century. However, its impact on material shaping and making tool life, and enhancement of their mechanical properties are quite remarkable. The selection of appropriate process parameters for CT is essential for cost reduction and optimum productivity. This study focuses on the influence of key parameters of CT cycles (i.e., soaking temperature and duration) on the friction and wear behavior of AISI H13 hot die steel under dry sliding conditions against hardened and tempered AISI D3 cold work tool steel (counter face) at varying sliding speeds and loads. Mathematical models have been developed for wear rate, the average coefficient of friction, and maximum contact temperature using the Box-Cox methodology. The developed mathematical models have been validated by comparing with the experimental results. Moreover, the optimum values of the process parameter have been employed to maximize the output and validate the same by confirmation of the experiments. To the best of our knowledge, this is the first study that demonstrates the modeling and optimization of sliding friction and wear characteristics of AISI H13 under varied CT cycles.

Keywords: cryogenics treatment; hot die steel; friction; wear; modeling; Box-Cox method; optimization

1 Introduction

Tools and dies used for the mass production of the material forming and shaping industries do not last forever because they wear out either by the steady growth of wear flats or by the accumulation of cracks, which lead to fractures. Consequently, the tool and die failure cause a sudden disruption, causing a delay in the production schedule. Apart from the lifetime of tools and dies, the replacement cost of worn tools (consumable cost) and the time to replace worn-out tools are significant in materials forming and shaping economics [1]. Indeed, a significant part of the improvement in the economic productivity of the components can be attributed to the use of long life tools and dies. Therefore, to further enhance efficiency and reduce costs in its operation, it is necessary to improve the tool and die materials used. AISI H13 is hot die steel (HDS) used for forging dies, inserts, punches, molds for die casting of aluminum, zinc, and magnesium. Nowadays, one of the latest techniques used in the industry to enhance the materials physical and mechanical properties is cryogenic treatment (CT). In contrast to the surface treatment in CT, the bulk properties of the materials as well as the surface properties are affected and its effects are permanent [2]. Moreover CT enhances the mechanical properties of tools and their life-span [3]. Several researchers

* Corresponding author: Sanjeev KATOCH, E-mail: katochsanjeev@gmail.com

List of symbols

B	Load (N)	HDS	Hot die steel
C	Sliding velocity (m/s)	SCT	Shallow cryogenic treatment
C1	Cryogenic treatment at −154 ˚C	ST	Soak time
C2	Cryogenic Treatment at −184 ˚C	T	Tempering
CHT	Conventional heat treatment	T_M	Maximum contact temperature (˚C)
CT	Cryogenic treatment	VFA	Vacuum furnace austenization
CWS	Cold work steel	WEDM	Wire electrical discharge machine
DCT	Deep cryogenic treatment	W_R	Wear rate (gm/m)
HBW 5/750	Brinell hardness using tungsten ball indenter diameter 5 mm at 750 kg Applied load	μ_a	Average coefficient of friction

have cited the worthiness of CT and its influence on the modification of the morphology of microstructures and enhancement in hardness of the CWS and HDS [4–7].

Tribological behavior of CWS 80CrMo12 5 was examined by Amini et al. [8] at different cryo-treatment holding times (i.e., 0, 6, 24, 48, 72, and 168 h); sliding velocity: 0.1, 0.15, and 0.2 m/s; load: 120 N and 160 N; up to 1,000 m. They concluded that 48 h soaking time at deep cryogenic temperature is optimum to achieve maximum wear resistance. Das et al. [9] demonstrated the influence of varied cryo-treatment on dry sliding wear behavior of AISI D2 CWS at different holding times (i.e., 0, 12, 36, 60, and 84 h) at different operating parameters of load and sliding velocity in a pin-on-disk wear test. They suggested that at DCT 36 h is optimum soaking time to achieve maximum wear resistance for AISI D2 CWS.

The wear resistance study of cryogenically treated (treatment temperature of −145 ˚C and soaking times of 12, 24, 36, 48, and 60 h) AISI 52100 bearing steel at a load of 10–20 N and sliding velocity of 0.15 m/s, using the ball-disk arrangement was carried out by Gunes et al. [10]. They claimed that 36 h was the optimal soaking time at the DCT to enable better wear resistance and also reported that delamination wear occurred with micro cracks in the DCT samples.

Yong et al. [11] studied the effect of deep cryogenic treatment on tungsten carbide tools for turning. They found that cryogenically treated tools have increased wear resistance to chipping while performing continuous cutting for short periods of time. Chipping times achieved were 1.3 times longer than non-cryogenically treated tools. Cryogenically treated tools lose their wear resistance when exposed to prolonged periods of high temperature during continuous cutting. Firouzdor et al. [12] demonstrated that deep CT enhances the wear resistance and tool life of the M2 HSS drill when used for dry high-speed drilling of normalized CK 40 steel. Precipitation of fine spherical carbides (spherical carbides are Fe3 M2, (M = (W, Mo, Cr, V) C) as a result of cryogenic treatment is responsible for wear resistance improvement. The main reason for improving fine carbide precipitation is due to super-saturation of martensite with decreasing temperature, leading to lattice distortion and thermodynamic instability of martensite; therefore, both carbon and alloy elements migrate to the nearby defects and segregate there which results in the formation of fine carbides on the subsequent warming up or tempering. As per this study, diffusion wear was found as the dominant wear mechanism**.** Huang et al. [13] analyzed the effect of CT on M2 tool steel microstructures using TEM and XRD. The results showed that CT can facilitate the formation of carbon clustering and increase the carbide density, thereby improving the wear resistance of steels. Precipitation of fine spherical carbides (spherical carbides are Fe3 M2, (M = (W, Mo, Cr, V) C) because CT is responsible for wear resistance improvement. Retained austenite eliminates and fine-dispersed eta (ή) carbide is precipitated.

It is evident from the literature review that due to the lack of visible changes on cryogenically treated

materials and the absence of organized fundamental research on the CT process, skepticism prevailed in the industry about this promising material treatment technique.

Though some researchers studied the effects of CT on the wear behavior of tool steel, but till now no research studies have been reported to develop and validate the models or to optimize the sliding friction and wear characteristics of HDS AISI H13 under varied CT cycles. The benefits of CT on the tool and die materials can only be exploited to its full potential if the CT cycles are optimized for the different grades of the tools and die steels according to the desired application. In this work, we attempt to develop the models of friction and wear behavior of varied cryogenically treated (soaking temperature and soaking duration) HDS AISI H13 sliding against hardened and tempered CWS AISI D3 (counter face) at varying sliding speeds and loads on a multi-tribotester under dry conditions. Hence, the authors claim the novelty of this study.

Mathematical modeling equations for wear rate (W_R), average coefficient of friction, and maximum contact temperature of varied cryogenically treated HDS AISI H13 are developed using the Box-Cox method of the response surface methodology (RSM) technique. To find the optimized conditions, the desirability function approach has been used to maximize the output parameter. In the sheet metal industry, HDS AISI H13 is used as a die for punching operations as well as for flash cutting in forging operations with CWS AISI D3 as the punch. This pair faces strong challenges regarding wear resistance in practical conditions. Hence, the HDS–CWS pair is selected for this experimental study. The counter-face material has higher hardness than the selected HDS.

2 Materials and methods

The AISI H13 material in the shape of round bars of diameter 16 mm and in the spheroidized annealed condition was used for this study. The chemical composition of the material was confirmed with the optical spark emission Spectrometer (DV6, Baird, USA), following ASTM E 415-2014 standards [14] and given in Table 1.

The hardness of the received material was checked with the calibrated optical Brinell hardness tester (B-3000(O), Saroj, India), using the Tungsten carbide ball of diameter 5 mm at 750 kg load with a dwell time of 15 s, as per the ASTM E 10-2008 standards [16]. The hardness of the received material is found to be in the range of 167–170 HBW 5/750. Blocks of 6.35 mm × 6.35 mm × 9 mm were machined from round bars using WEDM Make: Charmilles Tech. Switzerland in the longitudinal direction of received HDS AISI H13, as per the ASTM G 77-05 RA 2010 standards [17]. Vacuum heat treatment of machined samples was performed in the horizontal front loading electrically heated vacuum furnace (Hind High Vacuum, Bangalore, India, and capacity: 600 mm × 600 mm × 900 mm), at a vacuum level maintained to 10^{-2} mbar, austenization temperature of 1,040 ˚C, soak time at austenization temperature of 30 min, nitrogen gas quench to 27 ˚C at a gas pressure of 5 bar. Table 2 provides the details of the sample treatment condition and nomenclature. After the quenching samples were taken out from the furnace and divided into three groups, namely A3T: vacuum heat treated and three times tempered for 2 h, C1: vacuum heat treated plus cryogenic treated at −154 ˚C for varied soak times of 6, 21, and 36 h and tempered for 2 h, C2: vacuum heat treated plus cryogenic treated at −184 ˚C for varied soak times of 6, 21, and 36 h and tempered for 2 h. The A3T group samples of HDS H13 were tempered at 550, 570 and 620 ˚C respectively for 2 h, coded as H13 A3T, in a box-type electrically heated furnace, with a capacity of 300 kg, and a temperature range up to 750 ˚C. The C1 and C2 group samples of both grade materials were deeply, cryogenically treated at −154 ˚C and

Table 1 Chemical composition of HDS steel grade AISI H13 in weight% [15].

Element	%C	%Si	%Mn	%P	%S	%Ni	%Cr	%Mo	%V	%Co	%W
Observed value	0.39	0.88	0.32	0.018	0.007	0.04	5.00	1.27	0.93	0.01	0.18
Specified value	0.32 to 0.45	0.80 to 1.20	0.20 to 0.50	0.030Max	0.030Max	0.30Max	4.75 to 5.50	1.10 to 1.75	0.80 to 1.20	—	—

Table 2 Heat treatment sequence followed for HDS H13.

No.	Nomenclature	Depiction of treatment
1	A 3T	VFA = 1,040 ˚C, ST = 30 min, nitrogen gas quench, quench pressure = 5 bar, three *T* = 550 ˚C, 570 ˚C, 620 ˚C respectively for 2 h.
2	A T C1(6)T	VFA = 1,040 ˚C, ST = 30 min, nitrogen gas quench, quench pressure = 5 bar, *T* = 550 ˚C for 2 h, C1, ST = 6 h, *T* = 620 ˚C for 2 h.
3	AT C1(21)T	VFA = 1,040 ˚C, ST = 30 min, nitrogen gas quench, quench pressure = 5 bar, *T* = 550 ˚C for 2 h, C1, ST = 21 h, *T* = 620 ˚C for 2 h.
4	AT C1(36)T	VFA = 1,040 ˚C, ST = 30 min, nitrogen gas quench, quench pressure = 5 bar, *T* =550 ˚C for 2 h, C1, ST = 36 h, *T* = 620 ˚C for 2 h.
5	AT C2(6)T	VFA = 1,040 ˚C, ST = 30 min, nitrogen gas quench, quench pressure = 5 bar, *T* = 550 ˚C for 2 h, C2, ST = 6 h, *T* = 620 ˚C for 2 h.
6	AT C2(21)T	VFA = 1,040 ˚C, ST = 30 min, nitrogen gas quench, quench pressure = 5 bar, *T* =550 ˚C for 2 h, C2, ST = 21 h, *T* = 620 ˚C for 2 h.
7	AT C2(36)T	VFA = 1,040 ˚C, ST = 30 min, nitrogen gas quench, quench pressure = 5 bar, *T* = 550 ˚C for 2 h,C2, ST = 36 h, *T* = 620 ˚C for 2 h.

* Number in parentheses show the soaking times in h at the cryogenic temperature.

−184 ˚C, respectively, for varied soak times of 6, 21, and 36 h, with pre- and post-tempering at two different temperatures.

Varied CT was performed in a computer-controlled cryogenic processor with a tempering facility up to 150 ˚C (Make: Primero Enserve, Chennai, India). To maintain the stresses at a minimum level and to avoid thermal soaking of the material due to the abrupt temperature gradient of the case and core of the material, the slow cooling rate (1 ˚C/min) from ambient to deep CT soak temperature and a slow heating rate (1 ˚C/min) from cryo-treatment soak temperature to ambient temperature were selected.

A hardened CWS AISI D3 roller with the following dimensions: outer diameter = 60 mm, inner diameter = 25 mm, and thickness = 20 mm, was used as the counter-face material. This roller was hardened and tempered to a hardness of 52 HRc and its surface roughness was maintained at Ra<0.2 μm. The schematic of the wear test setup is shown in Fig. 1. Block-on-ring dry sliding wear tests were performed to assess the W_R, as per the ASTM standards designation G77-05 (Reapproved 2010) [17] using a multi-tribotester (Model: TR-30-M4, Make: DUCOM Instrument Pvt. Ltd. Bangalore, India). To attain the surface roughness in the range of 0.152–0.305 μm, faces of the blocks were grounded manually using silicon carbide emery paper of different grit sizes: 100, 220, 800, 1,000, 1,200, and 1,400 μm and finely polished

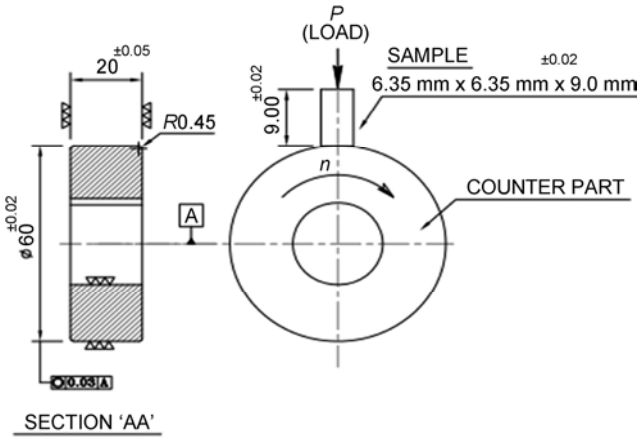


Fig. 1 Schematic presentation of the block on the ring wear test.

using a diamond slurry of grit size 6 μm with white kerosene oil as a suspension media on a rotating velvet cloth. Samples were cleaned in hexane prior to the wear test. The gap between the static block sample and the counter face of the roller was maintained as 2 mm for each experiment. After each test roller was polished using a silicon carbide emery paper of different grit sizes: 100, 400 and 1,000 μm they were then cleaned with hexane solution to maintain the average surface (Ra<0.2 μm). Before and after each experiment, the weight of the samples were measured using a precision electronic analytical balance (Model: HM-200, Make: A&D) with an accuracy of 0.00001 gm to calculate the weight loss. A non-contact type infrared thermometer (TFI20; Ebro; accuracy: ±2.5% or 2.5 ˚C;

resolution: 0.2 ˚C; measuring range: −30 ˚C to 180 ˚C) was used to measure the contact temperature at the junction of the sample and the counter face.

The experiments were carried out using normal load and sliding velocity in the range of 60–140 N and 0.628–1.885 m/s, respectively, in dry sliding conditions to measure the W_R and coefficient of friction. The W_R of each block was calculated from the weight loss, and each test was repeated three times for each condition to obtain the average value of weight loss. The duration of each experiment was 300 s. A comparison of the varied CT with respect to CHT has been made to identify its effects on the tribological behavior. Investigation of the tribological behavior in a range of parameters at various levels was carried out for the conventional and varied cryogenically treated samples at: (a) varying levels of cryogenic soak temperature (i.e., −155 ˚C and −184 ˚C), and (b) varying levels of cryogenic soak time (i.e., 6, 21, and 36 h). The wear test of conventional and varied cryogenically treated samples was performed using the full factorial design of the experiment. Levels of the input parameter for the wear test experiments are shown in Table 3. Machine capabilities were taken into consideration while selecting the test parameters.

A total of 175 experimental runs were conducted. The average of three replicate values of each run was taken as the dependent variable, or response, or yield (W_R, average coefficient of friction, and maximum contact temperature) to account for experimental errors. The experimental design was created with Design-Expert 7.1 (Stat-Ease, USA), using the full factorial design of experiments. This experiment design allows one to study the effect of each factor on the response variables, as well as their interaction with each other. To study the main and interaction effects of one categorical and two numeric factors on the response parameter, a statistical analysis of variance (ANOVA) was performed. To identify whether the factors are significant or not, an *F*-test in ANOVA was used, where higher *F*-values indicate that a factor has higher influence on the process.

Microhardness values (HV_1) were obtained for A3T, ATC1(6)T, ATC1(21)T, and ATC1(36)T AISI H13 HDS at five equally spaced points along the diagonal length of each specimen using the micro Vicker hardness tester (Model: MVK-H2, Make: Akashi), by following the ASTM E384-08a standards [18]. Table 4 depicts the results of the mean micro-hardness values along with their corresponding standard deviation and standard error of mean.

Figure 2 shows the FESEM micrograph photo of the

Table 3 Experimental variables and their different levels used for the block-on-ring dry sliding wear tests as input parameters.

Code	Factor	Unit	Levels values						
A	Treatment Type	—	A3T	ATC1(6)T	ATC1(21)T	ATC1(36)T	ATC2(6)T	ATC2(21)T	ATC2(36)T
	Treatment Temperature	˚C			−154			−184	
B	Load	N		60		80	100	120	140
C	Sliding Speed	m/s		0.628		0.942	1.257	1.571	1.885

Table 4 Mean micro-hardness of CHT and varied cryogenic treated samples.

Treatment Condition	Mean micro-hardness (HV_1)	Standard deviation	SE of mean
A3T	450.2	5.06	2.26
ATC1(6)T	464.4	4.61	2.06
ATC1(21)T	443.4	4.21	1.88
ATC1(36)T	433.6	3.57	1.6
ATC2(6)T	452.1	4.90	2.19
ATC2(21)T	434.8	6.37	2.85
ATC2(36)T	426.4	4.03	1.80

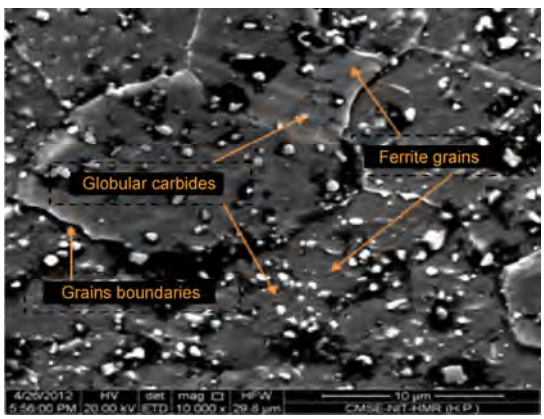


Fig. 2 FESEM micrograph photo of AISI H13 steel in the received condition.

material before the treatment. It depicts the globular carbides in the matrix of ferrite [19]. The microstructure of the received material indicates that the material is in an annealed condition which is necessary for the machining operation to prepare the samples.

The micrograph photo of the material after the treatment is shown in Fig. 3, which clearly shows the even distribution of the precipitated fine secondary globular carbide in the matrix of tempered martensite. Moreover, the results are in concurrence with previous studies [5, 9, 11] that reported the enhancement of numerous secondary carbides and the formation of homogenous carbides, which is responsible for the strengthening of the matrix, load bearing capacity, and wear resistance of tools and dies.

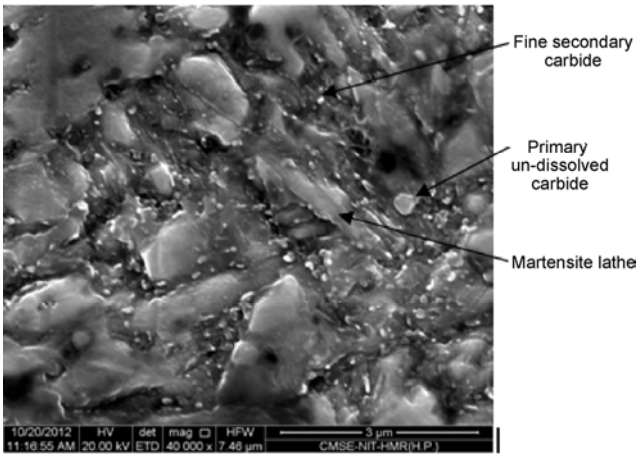


Fig. 3 FESEM image of the cryogenic treated sample of AISI-H13.

3 Results and discussion

The collection of experimental data for the dependent variable, or response (W_R, average coefficient friction, and rise in maximum contact temperature) permits the estimation of all main and interaction effects. Response surface methodology (RSM) was used to optimize the experiment variables that produce desirable values of the response. The results and discussion are divided between the following sections:

3.1 Evaluation of W_R at dry sliding conditions for AISI-H13

Figure 4 shows the normal plot of residuals distribution and illustrates that the residuals follow a normal distribution, as residuals follow a straight line except a few are scattered at the upper and lower ends. Diagnosis of the internally studentized residuals versus the experimental run order in the case of W_R is shown in Fig. 5. The studentized residuals method is commonly used to detect discrepant data and is defined as the residual divided by the estimated standard deviation of that residual. Figure 5 gives the residuals versus run order plot. It presents the random scatter in the range of ±3 standard deviations and all points are well fitted in the model. To determine the outlier data or values which are not easily predicted by the model, diagnosis of the response value as a function of the predicted response value in the case of W_R is performed.

Figure 6 represents the actual versus the predicted response values in the case of W_R at the dry sliding condition. It indicates the adequacy of the model over the range of data; as all the data points are scattered about the 45° diagonal reference line. The line just

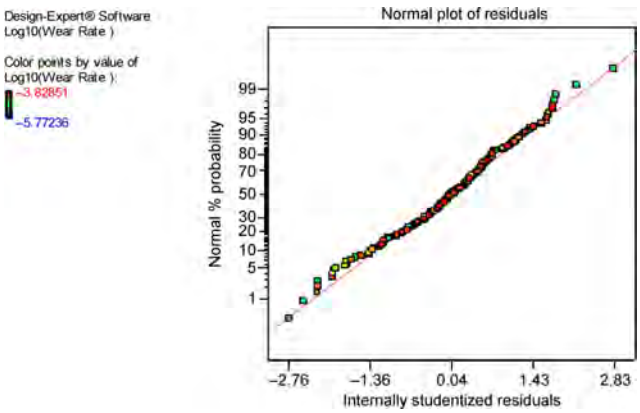


Fig. 4 Normal plot of residuals for W_R.

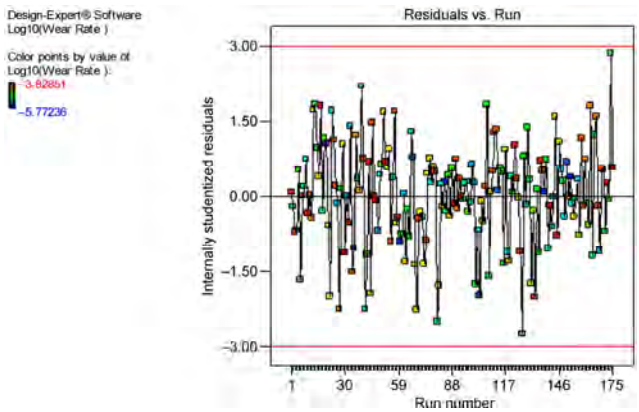


Fig. 5 Distribution of residuals over the experimentation run order for W_R.

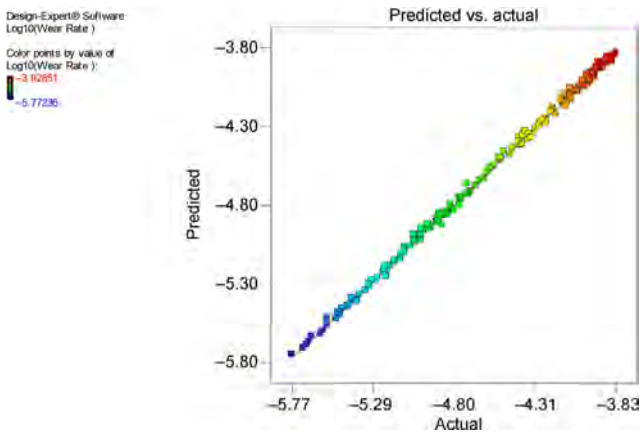


Fig. 6 Predicted versus actual response values for W_R.

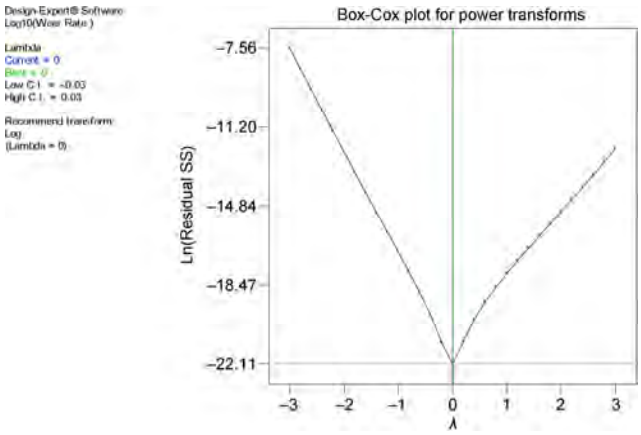


Fig. 7 Box-Cox plot for power transformation in the case of W_R.

passes through the middle of the data over a whole range of the data. In this case, the response of W_R to the input variables varies from the minimum value of 1.69×10^{-5} to 1.48×10^{-4}. The ratio of maximum to minimum in this case (87.87) is greater than the threshold limit of 10; this indicates that the transformation may be desirable. The normality of the process can always be improved by stabilizing the variance, which can be performed by the transformation.

To obtain a simple, normal, linear model that satisfies the aims of homogeneity of variance and normality data, the Box-Cox method was used to determine the power transformation for the dependent variable. Figure 7 shows the Box-Cox plot obtained using Design-Expert 7.1 software within ±3 standard deviations. The legend information on the left of Fig. 7 indicates that the minimum in Ln(Residual SS) occurs when $\lambda = 0.0$, the best λ value indicated in the plot by the long vertical (green line) at the center of the U-shaped curve. The point where the solid horizontal line cuts the U-shaped curve (identified by short pink lines) defines the upper and lower 95% confidence interval limit for the best λ value. The interval is $-0.03 < \lambda < 0.03$ and does not include $\lambda = 1$ (Fig. 7), thus ensuring that transformation will be helpful. The maximum likelihood estimate of λ is that value of λ which minimizes the error sum of squares of the fitted model. Design-Expert 7.1 recommends the log transformation ($\lambda = 0$) for the best fit of the model. The short vertical line (blue line) in Fig. 7 indicates the current transformation, which is the best-suggested transformation at $\lambda = 0$.

The selection of the model for the analysis is carried out using the data of coefficient of determination of the regression (R-squared). The R-squared values of data for different models obtained using Design-Expert 7.1 software are shown in Table 5. The "R-squared" values near to 1, always indicate that a regression line fits the data well. The suggested model for the data is cubic, with the "adjusted R-squared" value of 0.9986, "predicted R-Squared" value of 0.9981 and the p-value (Prob > F) < 0.0001.

After the elimination of all insignificant terms, ANOVA is modified manually by choosing the significant terms as shown in Table 6. In the new model, the overall F-test value is improved to 8,780.09 and remains significant with 13 degrees of freedom.

Therefore, the modified model has all the significant terms, i.e., A, B, C, BC, B^2, C^2, BC^2, and C^3. In the individual factors, the strongest influence on the W_R is found to be of the sliding velocity (with highest F value of 93,708.8), followed by load and treatment type. The significant influencing effect among the interaction is of BC only. The values of different tests under coefficient of regression analysis are determined and compared, as shown in Table 7. The "predicted R-Squared" value of 0.9983 is in reasonable agreement with the "adjusted R-Squared" value of 0.9985 in this model. Adequate precision, "Adeq precision" measures the signal to noise ratio. The reduced cubic model has an Adeq precision of 311.77, which is better than the full cubic model, and much higher than the minimum required value of four. The model can be used to navigate the design space, where the log of

Table 5 R-Squared values for different models in case of wear rate.

Source	Standard deviation	R-squared	Adjusted R-squared	Predicted R-squared	PRESS	Remarks
Linear	0.14	0.9428	0.94	0.9363	3.45	
2FI	0.12	0.9577	0.9518	0.945	2.97	
Quadratic	0.12	0.961	0.9551	0.9486	2.78	
Cubic	0.021	0.999	0.9986	0.9981	0.1	Suggested

Table 6 Modified ANOVA for the case of W_R.

Source	Sum of squares	DF	Mean Square	*F* value	p-value Prob > *F*	Remarks
Model	53.98	13	4.15	8780.09	< 0.0001	Significant
A-type of treatment	0.43	6	0.072	152.16	< 0.0001	Significant
B-load	6.22	1	6.22	13142.7	< 0.0001	Significant
C-sliding velocity	44.32	1	44.32	93708.7	< 0.0001	Significant
BC	0.8	1	0.8	1685.4	< 0.0001	Significant
B^2	0.11	1	0.11	239.04	< 0.0001	Significant
C^2	0.069	1	0.069	146.6	< 0.0001	Significant
BC^2	0.05	1	0.05	105.74	< 0.0001	Significant
C^3	1.99	1	1.99	4199.96	< 0.0001	Significant
Residual	0.076	161	4.73E–04			
Corrected total	54.06	174				

Table 7 Values of coefficient for full and modified cubic models from ANOVA analysis.

Full cubic model		Reduced cubic model	
Coefficient of regression	Value	Coefficient of regression	Value
Standard deviation	0.021	Standard deviation	0.022
Mean	–4.60	Mean	–4.60
C.V. %	0.45	C.V. %	0.47
PRESS	0.10	PRESS	0.090
R-squared	0.9990	R-squared	0.9986
Adjusted R-squared	0.9986	Adjusted R-squared	0.9985
Predicted R-squared	0.9981	Predicted R-squared	0.9983
Adeq precision	184.323	Adeq precision	311.677

W_R can be expressed by the following final regression equations:

Type of treatment: A3T

$$\text{Log}_{10}(W_R) = (-4.15 + 0.02 \times B - 7.84 \times C + 3.72E - 003 \times B \times C - 3.80E - 005 \times B^2 + 7.87 \times C^2 - 3.62E - 003 \times B \times C^2 - 2.02 \times C^3) \quad (1)$$

Type of treatment: ATC1(6)T

$$\text{Log}_{10}(W_R) = (-4.19 + 0.02 \times B - 7.84 \times C + 3.72E - 003 \times B \times C - 3.80E - 005 \times B^2 + 7.87 \times C^2 - 3.62E - 003 \times B \times C^2 - 2.02 \times C^3 \quad (2)$$

Type of treatment: ATC1(21)T

$$\text{Log}_{10}(W_R) = (-4.23 + 0.016 \times B - 7.84 \times C + 3.72E - 003 \times B \times C - 3.80E - 005 \times B^2 + 7.87 \times C^2 - 3.620E - 003 \times B \times C^2 - 2.02 \times C^3) \quad (3)$$

Type of treatment: ATC1(36)T

$$\text{Log}_{10}(W_R) = (-4.24 + 0.02 \times B - 7.84 \times C + 3.72261E - 003 \times B \times C - 3.80E - 005 \times B^2 + 7.87 \times C^2 - 3.62E - 003 \times B \times C^2 - 2.02 \times C^3) \quad (4)$$

Type of treatment: ATC2(6)T

$$\text{Log}_{10}(W_R) = (-4.27 + 0.01 \times B - 7.84 \times C + 3.72E - 003 \times$$

$B \times C - 3.80E - 005 \times B^2 + 7.87 \times C^2 - 3.62E - 003 \times B \times C^2 - 2.02 \times C^3)$ (5)

Type of treatment: ATC2(21)T

$\text{Log}_{10}(W_R) = (-4.31 + 0.02 \times B - 7.84 \times C + 3.72E - 003 \times B \times C - 3.80E - 005 \times B^2 + 7.87 \times C^2 - 3.621770E - 003 \times B \times C^2 - 2.02 \times C^3)$ (6)

Type of treatment: ATC2(36)T

$\text{Log}_{10}(W_R) = (-4.18 + 0.02 \times B - 7.84 \times C + 3.72E - 003 \times B \times C - 3.80E - 005 \times B^2 + 7.87 \times C^2 - 3.62E - 003 \times B \times C^2 - 2.02 \times C^3)$ (7)

The validation of Eqs. (1)–(7) is confirmed by comparing the trends of the experimental values of W_R with the corresponding theoretical values obtained from the ANOVA analysis of the model (Fig. 8) at varying sliding speeds and loads. Figure 8 shows the similar experimental and theoretical results with an average maximum error of 9% in the corresponding

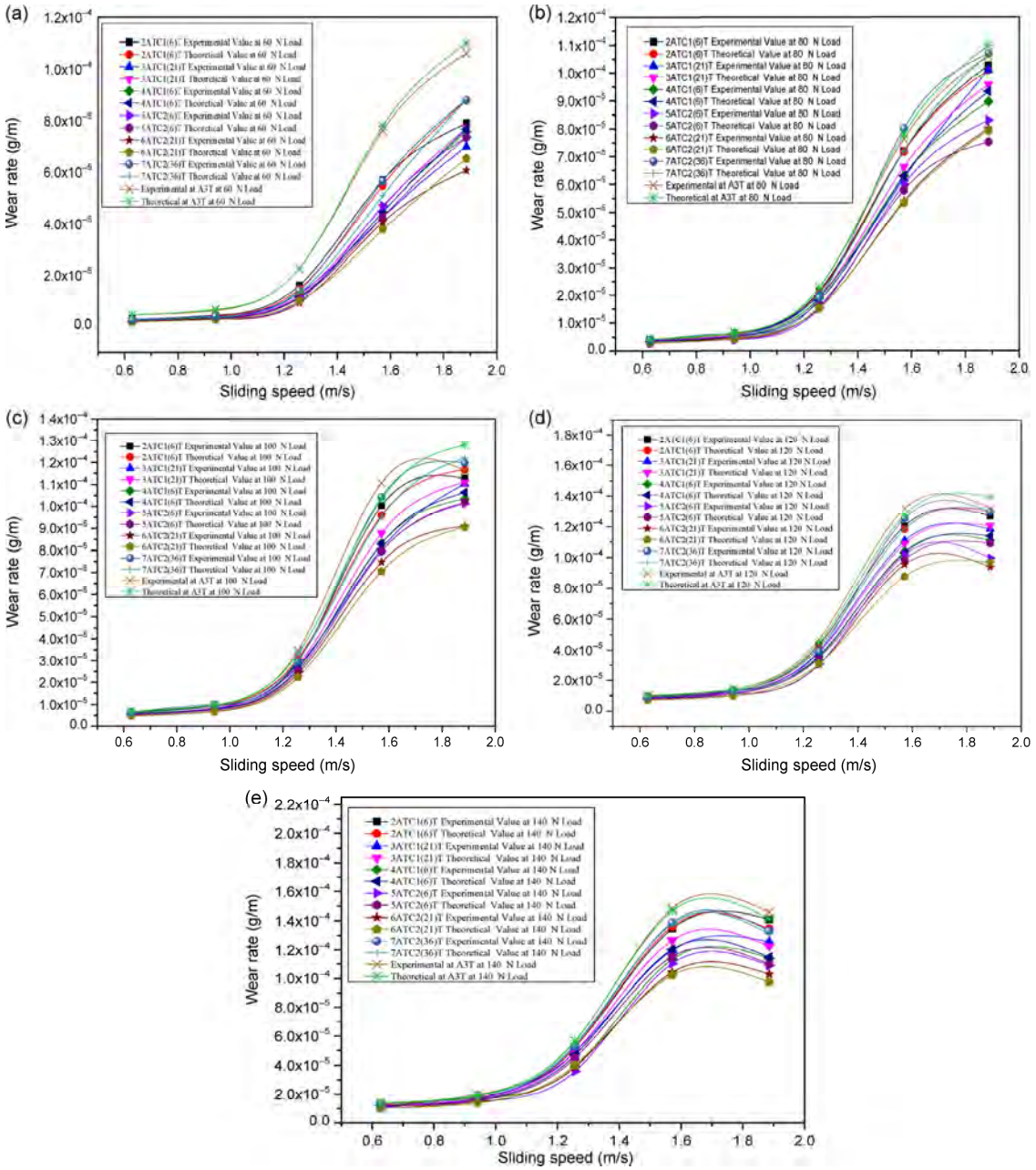


Fig. 8 Comparison of theoratical and experimental results for W_R at varying sliding speeds and loads: (a) 60 N, (b) 80 N, (c) 100 N, (d) 120 N, and (e) 140 N in the case of varied treatments: A3T, ATC1(6)T, ATC1(21)T, ATC1(36)T, ATC2(6)T, ATC2(21)T, and ATC2(36)T.

values. This confirms the validation of the model equations obtained for the W_R from this analysis and is thus acceptable.

It is observed that at a sliding speed of 0.628 m/s to 1.257 m/s, the W_R increases slowly, after that a sharp increase in W_R is evidenced at a sliding speed from 1.257 m/s to 1.571 m/s and it shows a decreasing trend at higher sliding speeds, i.e., 1.885 m/s for selected loads (60–140 N) for conventionally treated and all types of cryogenic treated material under investigation.

The worn surfaces are shown in Fig. 9. Worn surfaces depict the fractured ridges, surface cracks, and deformation lips as well as wear debris in the shape of large chunks of metalic plates for all the treatment conditions. On the other hand, the sizes of the debris metal plates vary with the treatment conditions and operating parameter, i.e., sliding velocity and applied normal load. The size of the wear debris found in the case of the CHT group samples is more compared to the cryogenic treated samples. In the case of the sample which was subjected to the 36 h soak time, at the CT temperature displayed surface cracks, fractured ridges, deformation lips, and wear debris in the shape of metal sheets, indicative of delamination wear [20]. The severe plastic deformation of the samples (ATC2 (36) T) due to the lower hardness is indicated in Table 4.

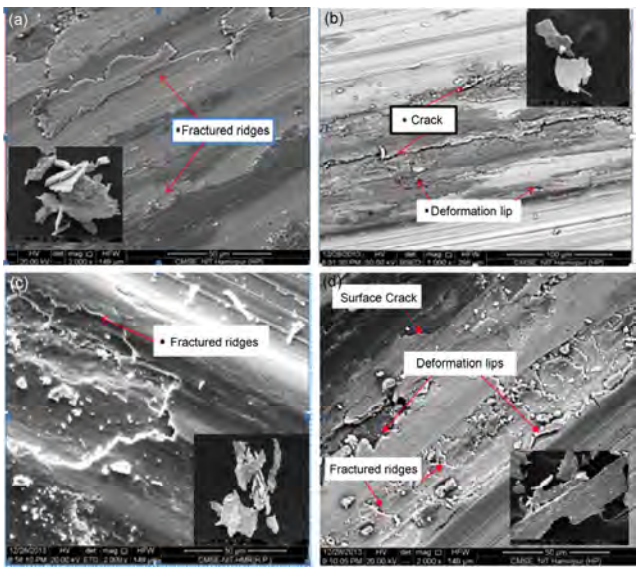


Fig. 9 FESEM micrograph photos of the worn surfaces generated under the wear test at a normal load of 140 N and sliding velocity of 1.885 m/s for different treatments: (a) A3T, (b) ATC1(6)T, (c) ATC1(21)T, and (d) ATC1(36)T. Insets represent the wear debris generated in the respective sample.

3.2 Evaluation of average coefficient of friction (μ_a) at dry sliding condition for AISI-H13

Figure 10 illustrates the analysis for the distribution of the residuals, and indicates that the residuals follow a normal distribution. Diagnosis of the internally studentized residuals versus the experimental run order in the case of the coefficient of friction, shown in Fig. 11, illustrates the random scatter in the range of ±3 standard deviations with all the points well fitted in the model. Thus, it proves the acceptability of the model and shows that the model satisfies the assumption for the ANOVA. The actual versus the predicted response values in the case of the average coefficient of friction are presented in Fig. 12, which depicts that the whole range of data points are scattered about the 45-degree diagonal reference line of the data. In this case, the response of the average coefficient of friction to the input variables varies from the minimum value of 0.17 to 0.22. The ratio of

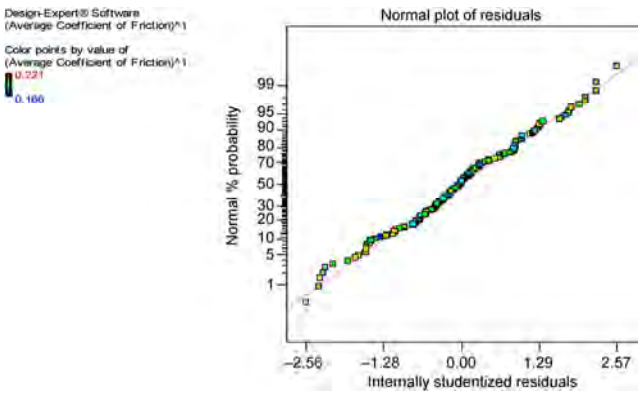


Fig. 10 Normal plot of residuals for the average coefficient of friction.

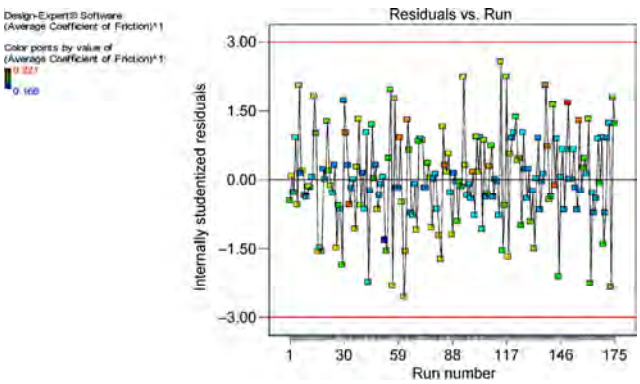


Fig. 11 Distribution of residuals over the experimentation run order for the average coefficient of friction.

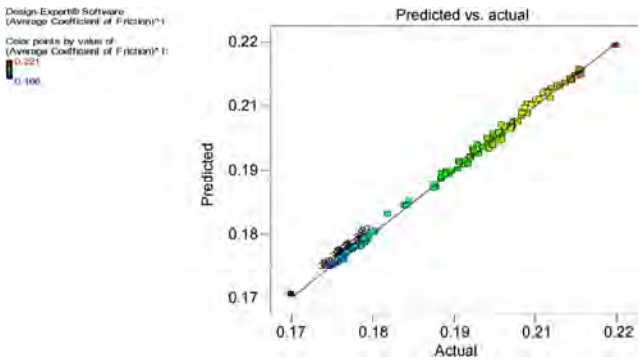


Fig. 12 Predicted versus actual response values for the average coefficient of friction.

maximum to minimum (1.33) is less than the threshold limit of 10, and thus it does not indicate the preference of transformation.

To further confirm this, the Box-Cox method is used to determine the power transformation for the dependent variable. Figure 13 shows the Box-Cox plot obtained using Design-Expert 7.1 software. The legend information on the left of Fig. 13 indicates that the minimum in Ln(Residual SS) occurs when $\lambda = -1.39$, the best λ value is indicated in the graph by the long vertical (green line) at the center of the U-shaped curve. The point where the solid horizontal line cuts the U-shaped curve (identified by the short pink lines) defines the upper and lower 95% confidence interval limit for the best λ value. The interval is $-2.43 < \lambda < -0.3$ and does not include $\lambda = 1$ (Fig. 13), thus ensuring that transformation will be helpful. Design-Expert 7.1 software recommends using the power transformation for the best fit of the model although one cannot rule out log, square root, inverse square root, or inverse transformation, as these all fall within the 95% confidence region. The current recommended transformation is at $\lambda = 1$ (indicated by blue line), which is a power transformation and is quite near to the best transformation at $\lambda = -1.39$.

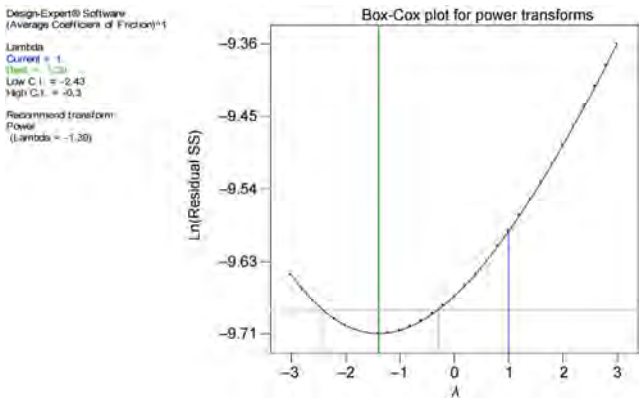


Fig. 13 Box-Cox plot for the average coefficient of friction.

Table 8 gives R-squared values of data for different models. The "R-squared" value near to one always indicates that a regression line fits the data well. The suggested model for the data is linear, with the "adjusted R-squared" value of 0.9974 and "predicted R-Squared" value of 0.9966, with the p-value (Prob > F) < 0.0001. After elimination of the insignificant terms, the modified ANOVA is shown in Table 9.

In the new model, the overall F-test value is improved to 1,547.724 and remains significant with 42 degrees of freedom. Therefore, the modified model has all the significant terms. In the individual factors, the strongest influence on the coefficient of friction is found to be due to the "A-type of treatment" (with the highest F value of 8,175), followed by load. The value of different tests under the coefficient of regression analysis was determined and compared as shown in Table 10. The "predicted R-squared" value of 0.9965 is in reasonable agreement with the "adjusted R-squared" value of 0.9973 in this model. The reduced cubic model has an Adeq precision of 149.70, which is better than the full linear model, and much higher than the minimum required value of four. The model can be used to navigate the design space, where the power

Table 8 R-Squared values for different models in the case of coefficient of friction.

Source	Standard deviation	R-squared	Adjusted R-squared	Predicted R-squared	PRESS	Remarks
Linear	4.44E−03	0.903	0.8983	0.8913	3.67E−03	
2FI	1.60E−03	0.9884	0.9869	0.9846	5.20E−04	
Quadratic	1.42E−03	0.991	0.9896	0.9876	4.19E−04	
Cubic	7.10E−04	0.9981	0.9974	0.9966	1.15E−04	Suggested

Table 9 Modified ANOVA for the coefficient of friction.

Source	Sum of squares	DF	Mean square	*F* value	p-value Prob > *F*	
Model	0.033706	42	0.000803	1547.724	<0.0001	Significant
A-type of treatment	0.025435	6	0.004239	8175.511	<0.0001	Significant
B-load	0.003596	1	0.003596	6935.531	<0.0001	Significant
C-sliding velocity	0.001468	1	0.001468	2831.016	<0.0001	Significant
AB	0.000828	6	0.000138	266.2975	<0.0001	Significant
AC	0.002028	6	0.000338	652.0205	<0.0001	Significant
BC	2.82E−05	1	2.82E−05	54.47658	<0.0001	Significant
B^2	6.59E−05	1	6.59E−05	127.0982	<0.0001	Significant
C^2	1.95E−05	1	1.95E−05	37.66527	<0.0001	Significant
ABC	3.89E−05	6	6.48E−06	12.5059	<0.0001	Significant
AB^2	0.000162	6	2.71E−05	52.20676	<0.0001	Significant
AC^2	2.59E−05	6	4.32E−06	8.325649	<0.0001	Significant
B^3	9.02E−06	1	9.02E−06	17.4038	<0.0001	Significant
Residual	6.84E−05	132	5.19E−07			
Corrected total	0.033774	174				

Table 10 Coefficients of regression for the full and modified cubic models from ANOVA analysis.

Full linear model		Reduced linear model	
Coefficient of regression	Value	Coefficient of regression	Value
Standard deviation	0.00071	Standard deviation	7.20E−04
Mean	0.190205	Mean	0.19
C.V. %	0.37331	C.V. %	0.38
PRESS	0.000115	PRESS	1.18E−04
R-squared	0.998074	R-squared	0.998
Adjusted R-squared	0.997403	Adjusted R-squared	0.9973
Predicted R-squared	0.996593	Predicted R-squared	0.9965
Adeq precision	147.8705	Adeq precision	149.705

of the coefficient of friction can be expressed by the following final regression equations:

Type of treatment: A3T

$$(\mu_a)^1 = (+0.25 - 6.15E - 004 \times B - 2.17E - 003 \times C - 4.27E - 005 \times B \times C + 4.55\,E - 006 \times B^2 - 7.082E - 003 \times C^2 - 1.67E - 008 \times B^3) \quad (8)$$

Type of treatment: ATC1(6)T

$$(\mu_a)^1 = (+0.25 - 9.37E - 004 \times B + 0.01 \times C - 9.29E - 005 \times B \times C + 6.84E - 006 \times B^2 - 3.49E - 003 \times C^2 - 1.67E - 008 \times B^3) \quad (9)$$

Type of treatment: ATC1(21)T

$$(\mu_a)^1 = (+0.25 - 1.03E - 003 \times B + 6.27E - 003 \times C - 7.46E - 005 \times B \times C + 7.19E - 006 \times B^2 - 1.86E - 003 \times C^2 - 1.67E - 008 \times B^3) \quad (10)$$

Type of treatment: ATC1(36)T

$$(\mu_a)1 = (+0.27 - 1.32E - 003 \times B + 4.95E - 004 \times C - 2.96E - 005 \times B \times C + 8.49E - 006 \times B^2 - 6.83E - 004 \times C^2 - 1.67E - 008 \times B^3) \quad (11)$$

Type of treatment:ATC2(6)T

$$(\mu_a)^1 = (+0.20 - 5.24E - 004 \times B - 7.60E - 004 \times C + 5.41E - 006 \times B \times C + 4.83E - 006 \times B^2 - 3.48E - 004 \times C^2 - 1.67E - 008 \times B^3) \quad (12)$$

Type of treatment: ATC2(21)T

$$(\mu_a)^1 = (+0.20 - 5.24E - 004 \times B - 7.61E - 004 \times C +$$

$5.41E-006 \times B \times C + 4.82E-006 \times B^2 - 3.48E-004 \times C^2 - 1.67E-008 \times B^3)$ (13)

Type of treatment: ATC2(36)T

$(\mu_a)1 = (+0.20 - 5.24E-004 \times B - 7.61E-004 \times C + 5.41E-006 \times B \times C + 4.83E-006 \times B^2 - 3.48E-004 \times C^2 - 1.67E-008 \times B^3)$ (14)

The validation of Eqs. (8)–(14) is confirmed by comparing the trends of experimental values of the average coefficient of friction in the wear test experiments with the corresponding theoretical values obtained from the ANOVA analysis of the model (Fig. 14) at varying sliding speeds and loads. The graphs demonstrate similar experimental and theoretical results with an average error of 3% in the corresponding values. This confirms the validation of the model equations obtained for the average coefficient of friction from this analysis and is thus acceptable.

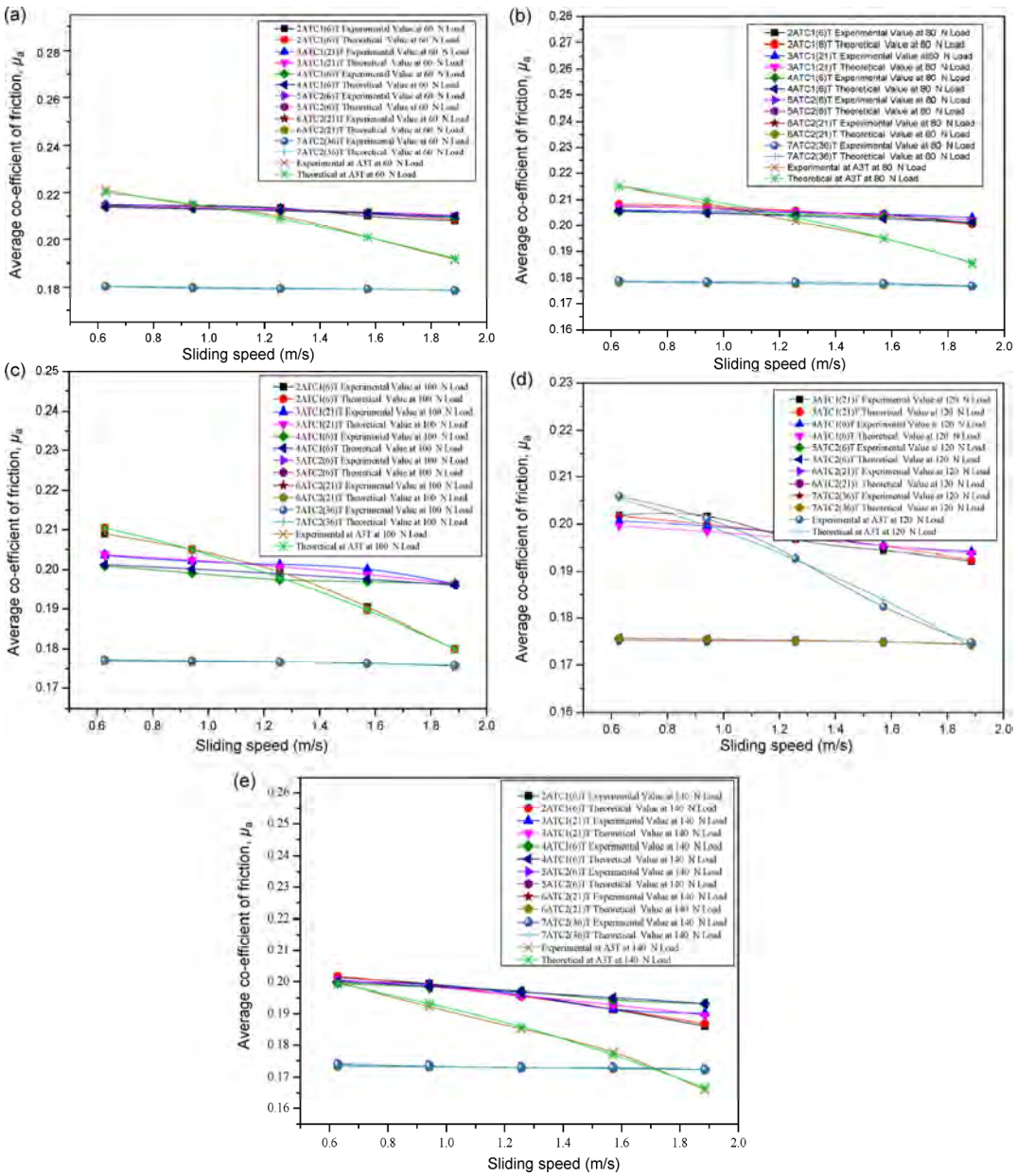


Fig. 14 Comparison of the theoretical and experimental results for the average coefficient of friction during the wear test at varying sliding speeds and loads: (a) 60 N, (b) 80 N, (c) 100 N, (d) 120 N, and (e) 140 N in the case of various treatments: A3T, ATC1(6)T, ATC1(21)T, ATC1(36)T, ATC2(6)T, ATC2(21)T, and ATC2(36)T.

3.3 Evaluation of the maximum contact temperature during the dry sliding condition for HDS AISI H13

Figure15 presents a distribution of the residuals for this analysis, and indicates that the residuals follow a normal distribution. Figure 16 illustrates the diagnosis of the internally studentized residuals versus the experimental run order for the case of contact temperature. This plot presents the random scatter in the range of ±3 standard deviations and all the points are well fitted in the model. The random scatter is found within the range of ±3 standard deviation for internally studentized residuals. The process has the similar trend about the mean line and is stable, which proves the acceptability of the model and shows that the model satisfies the assumption for the ANOVA. Figure 17 represents the actual versus the predicted response values for the case of contact temperature. It indicates the adequacy of the model over the range of data; as all the data points are scattered about the 45 degree diagonal reference line.

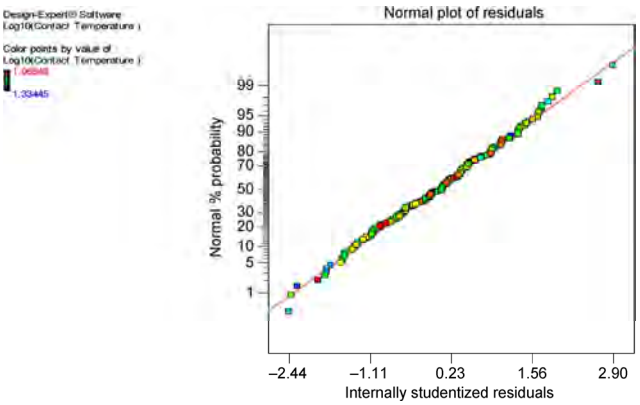


Fig. 15 Normal plot of residuals for maximum contact temperature.

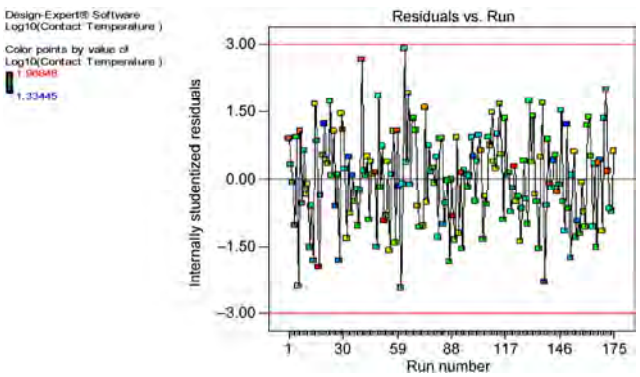


Fig. 16 Distribution of residuals over the experimentation run order for maximum contact temperature.

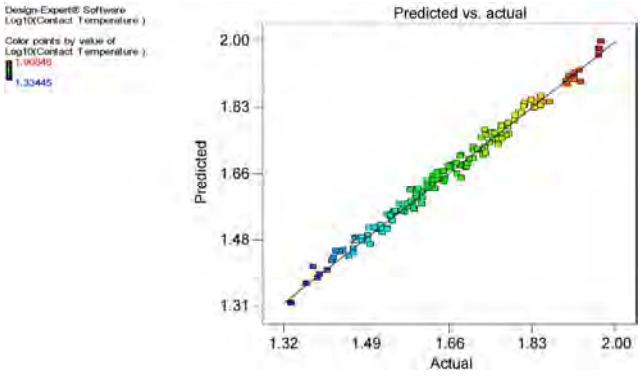


Fig. 17 Predicted as a function of actual response values for maximum contact temperature.

In this case, the response of contact temperature to the input variables varies from the minimum value of 21.6 to 93.0. The ratio of maximum to minimum is 4.30, which is well within the threshold limit of 10, therefore, does not indicate the preference of transformation. To further confirm this, the Box-Cox method was used to determine the power transformation for the dependent variable. Figure 18 shows the Box-Cox plot within ±3 standard deviations. The legend information on the left of Fig. 18 indicates that the minimum in Ln(Residual SS) occurs when $\lambda = -0.08$, the best λ value indicated in the plot by the long vertical (green line) at the center of the U-shaped curve. The point where the solid horizontal line cuts the U-shaped curve (identified by short pink lines) defines the upper and lower 95% confidence interval limit for the best λ value. The interval is $-0.36 < \lambda < 0.19$ and does not include the $\lambda = 1$ (Fig. 18), thus ensuring that the transformation will be helpful. Design-Expert 7.1 software recommends the log transformation ($\lambda = 0$)

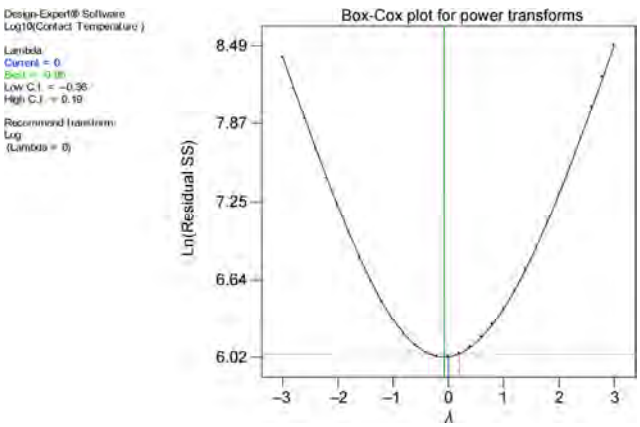


Fig. 18 Box-Cox plot for power transformation in the case of maximum contact temperature.

for the best fit of the model although one cannot rule out linear, cubic, quadratic and square root, transformation, as all these fall within the 95% confidence region. The short vertical blue line in Fig. 18 at $\lambda = 0$ indicates the current transformation, which is quite near to the best-suggested transformation at $\lambda = -0.08$. The R-squared values of data for different models obtained using Design-Expert 7.1 software are shown in Table 11. The "R-squared" value near to one always indicates that a regression line fits the data well. The suggested model for the data is cubic, with the "adjusted R-squared" value of 0.9905 and "predicted R-squared" value of 0.9840, with the p-value (Prob > *F*) < 0.0001. To obtain a more desirable model, all insignificant terms are eliminated and the ANOVA is modified accordingly, as shown in Table 12. In the new model, the overall F-test value is improved to 481.2 and remains significant with 30 degrees of freedom. Therefore, the modified model has all the significant terms, i.e., *A*, *B*, *C*, *AB*, *AC*, *BC*, BC^2, and AB^2. In the individual factors, the strongest influence on the contact temperature is found to be due to *C*: sliding velocity (with the highest *F* value of 9,760.87), followed by load and treatment type. The next significant influencing effects among the interaction are due to *B*: load, *A*: type of treatment, and *AB*. In the second order term, the strongest influence is found to be that of BC^2.

The values for different tests under the coefficient of regression analysis are determined and compared as shown in Table 13. The "predicted R-squared" value of 0.9844 is in reasonable agreement with the "adjusted R-squared" value of 0.9881 in this model. The reduced quadratic model has an Adeq precision of 98.502, which is better than the full cubic model, and much higher than the minimum required value of four. The model can be used to navigate the design space, where the log of contact temperature can be expressed by the

Table 11 R-Squared values for different models in case of maximum contact temperature.

Source	Standard deviation	R-squared	Adjusted R-squared	Predicted R-squared	PRESS	
Linear	7.30E−03	0.9325	0.9292	0.9245	9.89E−03	
2FI	5.37E−03	0.9663	0.9617	0.9532	6.13E−03	
Quadratic	4.66E−03	0.975	0.9712	0.9645	4.65E−03	
Cubic	3.30E−03	0.9893	0.9856	0.9789	2.76E−03	Suggested

Table 12 Modified ANOVA for maximum contact temperature.

Source	Sum of Squares	DF	Mean Square	*F* Value	p-value Prob > *F*	
Model	3.85	30	0.13	481.2	<0.0001	Significant
A-type of treatment	0.062	6	0.01	38.43	<0.0001	Significant
B-load	1.56	1	1.56	5827	<0.0001	Significant
C-sliding velocity	2.07	1	2.07	7751.78	<0.0001	Significant
AB	0.11	6	0.019	69.9	<0.0001	Significant
AC	0.018	6	3.05E−03	11.42	<0.0001	Significant
BC	0.013	1	0.013	50.19	<0.0001	Significant
B^2	4.95E−04	1	4.95E−04	1.85	0.1755	
C^2	2.09E−03	1	2.09E−03	7.81	0.0059	Significant
AB^2	0.014	6	2.25E−03	8.44	<0.0001	Significant
BC^2	7.50E−03	1	7.50E−03	28.09	<0.0001	Significant
Residual	0.038	144	2.67E−04			
Corrected total	3.89	174				

Table 13 Values of coefficient for full and modified quadratic models from ANOVA analysis.

Full linear model		Reduced linear model	
Coefficient of regression	Value	Coefficient of regression	Value
Standard deviation	0.015	Standard deviation	0.016
Mean	1.65	Mean	1.65
C.V. %	0.88	C.V. %	0.99
PRESS	0.062	PRESS	0.061
R-squared	0.9930	R-squared	0.9901
Adjusted R-squared	0.9905	Adjusted R-squared	0.9881
Predicted R-squared	0.9840	PredictedR-squared	0.9844
Adeq precision	89.217	Adeq precision	98.502

following final regression equations:

Type of treatment: A3T

$$\mathrm{Log}_{10}(T_M) = (+1.04 + 3.43 - 003 \times B - 0.10 \times C + 4.22E - 003 \times B \times C - 9.43E - 006 \times B^2 + 0.12 \times C^2 - 1.40E - 003 \times B \times C^2) \quad (15)$$

Type of treatment: ATC1(6)T

$$\mathrm{Log}_{10}(T_M) = (+1.20 + 7.52E - 004 \times B - 0.09 \times C + 4.22E - 003 \times B \times C + 1.04E - 006 \times B^2 + 0.12 \times C^2 - 1.40E - 003 \times B \times C^2) \quad (16)$$

Type of treatment: ATC1(21)T

$$\mathrm{Log}_{10}(T_M) = (+1.14 + 4.29E - 003 \times B - 0.16 \times C + 4.22E - 003 \times B \times C - 1.66E - 005 \times B^2 + 0.12 \times C^2 - 1.40E - 003 \times B \times C^2) \quad (17)$$

Type of treatment: ATC1(36)T

$$\mathrm{Log}_{10}(T_M) = (+1.37 + 2.59E - 004 \times B - 0.13 \times C + 4.22E - 003 \times B \times C - 5.01E - 006 \times B^2 + 0.12 \times C^2 - 1.40E - 003 \times B \times C^2) \quad (18)$$

Type of treatment: ATC2(6)T

$$\mathrm{Log}_{10}(T_M) = (+1.64 - 6.01E - 003 \times B - 0.15 \times C + 4.22E - 003 \times B \times C + 2.73E - 005 \times B^2 + 0.12 \times C^2 - 1.40E - 003 \times B \times C^2) \quad (19)$$

Type of treatment: ATC2(21)T

$$\mathrm{Log}_{10}(T_M) = (+1.26 + 1.07E - 003 \times B - 0.13 \times C + 4.22E - 003 \times B \times C - 5.65E - 006 \times B^2 + 0.12 \times C^2 - 1.40E - 003 \times B \times C^2) \quad (20)$$

Type of treatment: ATC2(36)T

$$\mathrm{Log}_{10}(T_M) = (+1.04 + 3.40E - 003 \times B - 0.10 \times C + 4.227E - 003 \times B \times C - 9.30E - 006 \times B^2 + 0.12 \times C^2 - 1.40E - 003 \times B \times C^2) \quad (21)$$

The validation of Eqs. (15)–(21) is confirmed by comparing the trends of experimental values of temperature at the junction of the block sample with the ring and with the corresponding theoratical values obtained from the ANOVA analysis of the (Fig. 19) at varying sliding speeds and loads. The graphs show similar experimental and theoratical results, with an average error of 8% in the corresponding values. This confirms the validation of the model equations obtained for contact temperature from this analysis and is thus acceptable.

It is observed that as the sliding velocity varies from 0.628 m/s to 1.885 m/s, the maximum contact temperature increases (from 21.6 ˚C to 93.0 ˚C) for selected loads (60–140 N) for the conventionally treated and all types of cryogenically treated material under investigation. As the temperature increases, the W_R decreases.

The effects of macroscopic temperature on wear behavior of friction pairs were evaluated using the relationship of sliding velocity and Peclet number (Pe). The Peclet number (Pe) for this study ranges from 0.86 to 7.7, which is less than 10. As the sliding velocity, or Pe increases, the maximum surface contact temperature decreases due to the transfer of heat to the moving mass.

While the front edge of the moving mass approaches the heat supply source, it is at a lower surface temperature in comparison to the heat source. The thermal diffusivity and heat capacity of the material is finite. Hence, to take in heat and to increase its temperature, this requires time. With the increase in moving mass

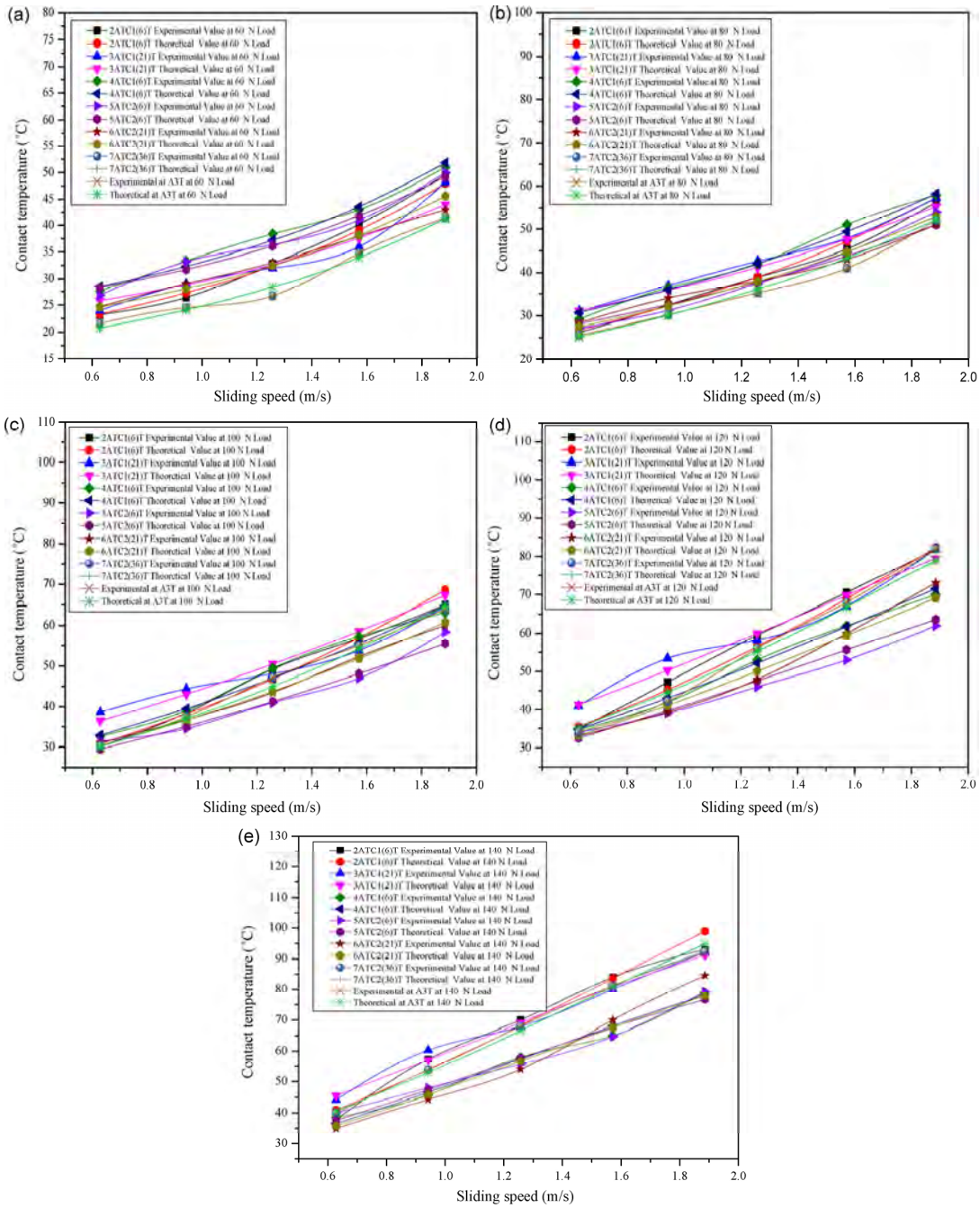


Fig. 19 Comparison of the theoretical and experimental results for the rise in contact temperature during the wear test at varying sliding speeds and loads: (a) 60 N, (b) 80 N, (c) 100 N, (d) 120 N, and (e) 140 N in the case of various treatments: A3T, ATC1(6)T, ATC1(21)T, ATC1(36)T, ATC2(6)T, ATC2(21)T, and ATC2(36)T.

velocity, it spends less time underneath the heat source for a specific volume of material. Thus, the increase in temperature will be less [21].

4 Optimization

The desirability function approach of the RSM technique has been used in this study to determine the optimum parameters to achieve a low value of W_R, average coefficient of friction, and maximum contact temperature. This study has three different responses for the variables, each of which is modeled with a different polynomial equation. The model for W_R is a natural log, the model for the average coefficient

of friction is linear and the model for the contact temperature is a natural log. Accordingly, each will have a different graph with different optimal factor settings. The values of the variables with maximum total desirability are considered to be optimal parameter conditions. The ranges and goals of each input process parameter and measured response (optimization criterion) are provided in Table 14, which is used to minimize the rate of wear, average coefficient of friction, and contact temperature, with the highest desirability function for the optimized setting for the desired output response.

The default value of "1" for the weight is assigned to the variables to achieve the maximum desirability function by giving equal weight to all goals. Considering the importance of each variable value given in Table 14, a set of 24 optimal solutions with the highest desirability function are obtained within the specifically designed space constraints. Table 14 illustrates the optimized results for minimizing the tribological parameters by considering seven combinations of categorical factor levels and the results are summarized in Table 15.

Further, conformity tests are carried out under the same conditions to validate the adequacy of the mathematical correlations. The objective of the confirmation experiments is to demonstrate the validity of the mathematical model derived from a designed experiment. The average of the results from the confirmation experiment is compared with the predicted average based on the parameters and levels tested, as shown in Table 16, which illustrates that the experimental values are in reasonable agreement with the predicted response values. The variations between the experimental and the predicted results are of the order of 2%–9%.

3D plots for desirability are constructed to consider the different possible combinations of the numeric variables as load and sliding velocity at the abscissa

Table 14 Constraints for determining the optimum values of the tribological parameters.

Parameter	Goal	Weight	Importance
Type of treatment (A)	A3T $\leqslant A \leqslant$ ATC(36)T	1	3
Load (B), N	$60 \leqslant B \leqslant 140$	1	3
Sliding velocity (C), m/s	$0.628 \leqslant C \leqslant 1.885$	1	3
Wear rate, W_R (D), gm/m	$1.69 \times 10^{-5} \leqslant D \leqslant 1.48 \times 10^{-4}$	1	5
Average coefficient of friction (E)	$0.16 \leqslant E \leqslant 0.22$	1	3
Contact temperature (F), °C	$21.6 \leqslant F \leqslant 93$	1	3

Table 15 Optimization results for the tribological parameters.

S. No.	Type of treatment, (A)	Load (B), N	Sliding velocity (C), m/s	W_R (D), g/m	Average coefficient of friction, (E)	Contact temperature (F), °C	Desirability
1	ATC2(21)T	60	0.687	1.73E−06	0.18	26.46	0.89
2	ATC2(21)T	60	0.682	1.73E−06	0.18	26.41	0.89
3	ATC2(21)T	60	0.696	1.73E−06	0.18	26.54	0.89
4	ATC2(21)T	60	0.703	1.74E−06	0.18	26.61	0.89
5	ATC2(21)T	60	0.661	1.74E−06	0.18	26.23	0.89

Table 16 Results of the conformity tests for the tribological parameters.

Test No.	Process parameter			Response parameter (predicted value)			Response parameter (experimental value)			% Error		
	Treatment type	B (N)	C (m/s)	W_R (gm/m)	μ_a	T_M (°C)	W_R gm/m	μ_a	T_M (°C)	W_R (gm/m)	μ_a	T_M (°C)
1	ATC2(21)T	60	0.687	1.73E−06	0.18	26.46	1.88E−06	0.18	24.5	8.67	2.08	7.41
2	ATC2(21)T	60	0.682	1.73E−06	0.18	26.41	1.83E−06	0.19	27.4	5.78	2.63	3.73
3	ATC2(21)T	60	0.696	1.73E−06	0.18	26.54	1.80E−06	0.19	24.9	4.04	3.19	6.18

(Fig. 20(a)), keeping all the response variables at the minimum to achieve the maximum possible life. The chosen categorical factor is ATC2(21)T following the result as shown in Table 15. Respective contour plots of desirability as a function of sliding velocity and loads are shown in Fig. 20(b). It can be interpreted from Figs. 20(a) and 20(b) that the overall desirability value is less in the region of the higher value of the load and sliding velocity. The desirability value is higher at the lower load and sliding velocity. The contour plot of Fig. 20(b) shows the sensitivity of the results to the condition. The optimal region was located in the lower left-side of the plot closer to the smaller values of the applied load and sliding velocity, represented by the maximum value of the predicted desirability (red colored region), which is 0.89. This region displayed overall the greatest desirability of 0.89 at the center. The change in color of the region while moving away from the highest desirability point indicates a gradual reduction of the desirability. The sensitivity of the response is also indicated by the shape of the contour lines in Fig. 20(b).

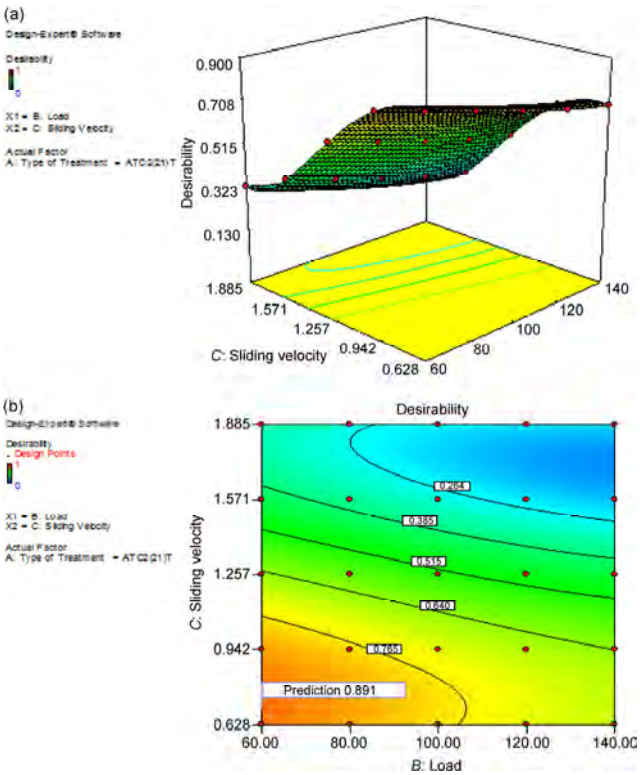


Fig. 20 (a) 3D plot, (b) contour plot for desirability versus sliding velocity and load.

5 Conclusions

In this study, mathematical models have been proposed to predict the tribological behavior (W_R, the average coefficient of friction (μ_a), and the maximum contact temperature (T_M)) of varied cryogenically treated HDS H13 against cold-work steel AISI D3 under a different set of operating parameters. Furthermore, optimal conditions are identified for the minimum W_R. The major conclusions drawn from this study are:

The present study demonstrates that the soaking time of 21 h at −184 ˚C for HDS H13, is the optimum soak time to have the maximum wear resistance.

The parameter sliding velocity (C) influences the W_R more in comparison to the second parameter load (B). The first parameter of the study (type of treatment) treatment ATC2(21)T gives the optimal set of conditions possessing the highest desirability value (0.891) and is selected for the desired response, which is: load: 60 N; sliding velocity: 0.687 m/s; W_R: 1.73×10^{-4} gm/m; average coefficient of friction (μ_a): 0.18; and maximum contact temperature: 26.46 ˚C.

The CT enhances the wear resistance of HDS H13 in comparison to the CHT, up to the holding time of 21 h at −184 ˚C, beyond this it shows a decrease with further increase in the soak time.

The morphology of worn surfaces of cryogenically treated samples changes from mild to severe as the sliding velocity and applied load increases in the chosen set of parameters for the experiment.

Wear debris has the shape of a plate-form of metal and appeared to delaminate from the samples surface due to sub-surface cracks and plastic deformation.

An average error of 9%, 3%, and 8% exists in the experimental and theoretical results obtained using the model equations for W_R, the average coefficient of friction (μ_a) and the maximum contact temperature (T_M) respectively.

The experimental values are in reasonable agreement with the predicted response values in the case of optimization. The variations between the experimental and the predicted results are of the order of 8.7%, 3.2%, and 7.4% for the W_R, the average coefficient of friction, and the maximum contact temperature, respectively, for the selected optimized treatment type ATC2(21)T.

Acknowledgement

The authors gratefully acknowledge their gratitude to the National Institute of Technology, Hamirpur to grant funds for the procurement of the hot die steel material, making available their tribological test; Institute of Auto Parts and Hand tools Technology, Ludhiana for extending facilities under the expert supervision of spectroscopic analysis, vacuum heat treatment and cryogenic treatment facility required for the study.

References

[1] Roberts G. *Tools Steels*. 5th Ed. Materials Park, OH, USA: ASM International, 1998: 46–66.

[2] Lal D M, Renganarayanan S, Kalanidhi A. Cryogenic treatment to augment wear resistance of tool and die Steel. *Cryogenics* **41**: 149–55 (2001)

[3] Barron R F. Effects of cryogenic treatment on lath tool wear. *Progress in Refrigeration Science and Technology, Publishing Co. Westport* **1**: 529–34 (1973)

[4] Das D, Dutta A K, Ray K K. Sub-Zero treatments of AISI D2 Steel: Part I. Microstructure and hardness. *Mater Sci Eng A* **527**: 2182–2193 (2010)

[5] Mehtedi M E, Ricci P, Drudi L, Mohtadi S E, Cabibbo M, Spigarelli S. Analysis of the effect of deep cryogenic treatment on the hardness and microstructure of X30CrMoN 15 1 steel. *Mater Design* **33**: 136–144 (2012)

[6] Koneshlou M, Meshinchi K, Khomamizadeh F. Effect of cryogenic treatment on microstructure, mechanical and wear behaviors of AISI H13 hot work tool steel. *Cryogenics* **51**: 55–61 (2011)

[7] Katoch S, Sehgal R, Singh V. Effect of cryogenic treatment on hardness, microstructure and wear behavior of hot die steel grade AISI-H13. In *Proceeding of International Conference on Advances in Tribology and Engineering Systems*, 2014: 159–166.

[8] Amini K, Nategh S, Shafyei A. Influence of different cryo-treatments on tribological behavior of 80CrMo12 5 cold work tool steel. *Mater Design* **31**: 4666–4675 (2010)

[9] Das D, Dutta A K, Ray K K. Influence of varied cryo-treatment on the wear behavior of AISI D2 steel. *Wear* **266**: 297–309 (2009)

[10] Gunes I, Cicek A, Aslantas K, Kara F. Effect of deep cryogenic treatment on wear resistance of AISI 52100 bearing steel. *Trans Indian Inst Met* **67**: 909–917 (2014)

[11] Yong A Y L, Seah K H W, Rehman M. Performance of cryogenically treated tungsten carbide tools in milling operations. *Int J Adv Manuf Technol* **32**: 638–643 (2007)

[12] Firouzdor V, Nejati E, Khomamizadeh F. Effect of deep cryogenic treatment on wear resistance and tool life of M2 HSS drill. *J Mater Process Technol* **206**: 467–472 (2008)

[13] Huang Y, Zhu Y T, Liao X Z, Beyerlein I J, Bourlce M A, Mitchell T E, Microstructure of cryogenic treated M2 Tool Steel. *Mater Sci Eng A* **339**: 241–244 (2003)

[14] ASTM E 415-2014. Standard test method for Analysis of carbon and low alloy steel by spark atomic emission spectrometry. *ASTM Annual Book of Standards*. West Conshohocken, PA, United States.

[15] Bayer A M, Vasco T, Walton L R. Properties and selection: Iron, steels and high performance alloys. *ASM Handbook, 3rd* ed. ASM International, Materials Park, OH, USA, 1995.

[16] ASTM E10-2010. Standard test method for Brinell hardness of metallic materials. *ASTM Annual Book of Standards.* West Conshohocken, PA, United States, 2010: 91–122.

[17] ASTM G77-2010. Standard test method for ranking resistance of materials to sliding wear using block-on-ring wear test. *ASTM Annual Book of Standards.* West Conshohocken, PA, United States.

[18] ASTM E384-08a (2009). Standard test method for micro indentation hardness of materials. *ASTM Annual Book of Standard, Vol. 3.01*. West Conshohocken, PA, United States.

[19] George F Vander Voort. Metallography and microstructure. *ASM Handbook, 9th ed*. ASM International, Materials Park, OH, USA, 1995: 257−263.

[20] Suh N P. An overview of the delamination theory of wear. *Wear* **44**:1–16 (1977)

[21] Sharma M D, Sehgal R, Pant M. Tribological behavior of Ti3Al2.5V alloy against EN-31 steel under dry condition. *Tribol Trans* **59**(3): 451–461 (2016)

Sanjeev KATOCH. He was graduated in science from Government Degree Collage Hamirpur, Himachal Pradesh University, Shimla, India in 1993. He obtained his master degree in materials science from Thapar University, Patiala, Punjab, India in 1997. He is currently pursuing his Ph.D. degree in the Center for Materials Science & Engineering, National Institute of Technology, Hamirpur, Himachal Pradesh, India. Along with his research work, he is working as deputy manager (Testing & Evaluation and Heat treatment Department) with Institute for Auto Parts & Hand Tools Technology, Ludhiana, Punjab, India. He has 18 years of experience in the heat treatment of materials, testing, evaluation of auto component & tool material through metallurgical, and mechanical & NDT techniques. His area of research/ interest is cryogenic processing of materials, tribological behavior of tools and dies, and failure analysis and non-destructive evaluation of materials and components. He has several publications in various reputed journals and international conference proceedings.

Rakesh SEHGAL. He is a professor in the Department of Mechanical Engineering at National Institute of Technology (Institute of higher learning-Government of India), Hamirpur, Himachal Pradesh, India. He obtained his M.Tech. degree in design of mechanical equipment from Indian Institute of Technology, Delhi, India, and received his PhD degree in mechanical engineering from Kurukshetra University, Kurukshetra, India in 2001. Dr. Sehgal has also completed his Post-Doctorate under PDF Award scheme of University Grants Commission, Ministry of Human Resource Development, Govt. of India. His research interest includes tribology, reliability modeling, and machining and optimization of production systems. He has two years of field and 25 years of teaching and research experience. He has published in journals such as *Transactions of the ASME, Journal of Tribology, Tribology International, Tribology Online, Proc. IMechE Part B: Journal of Engineering Manufacture, International Journal for Manufacturing Science & Production, International Journal of Precision and Manufacturing, International Journal of Production Economics, Journal of Reliability & Systems Safety, Journal of Intelligent Manufacturing, Materials and Manufacturing Processes, Applied Acoustics, Journal of Vibration and Acoustics*, among others.

Vishal SINGH. He is presently an associate professor at the Centre for Materials Science and Engineering, National Institute of Technology (Institute of higher learning-Government of India), Hamirpur, Himachal Pradesh, India. He received a B.E. degree in metallurgical engineering from University of Roorkee, Roorkee, India in 1992, an M. Tech. degree in metallurgical engineering from Indian Institute of Technology (IIT), Bombay, India in 1995 and a Ph.D. degree from Indian Institute of Technology (IIT), Bombay, India in 2003. His research interests include electrical, mechanical, and tribological behavior of polymer composites/ nano-composites, physical metallurgy, and ceramic processing and characterization. He has more than twelve years of research and teaching experience in materials science related areas and his research work has been published in several international journals and conference proceedings of repute.

Friction **5**(1): 87–98 (2017)
DOI 10.1007/s40544-017-0140-3
ISSN 2223-7690
CN 10-1237/TH
RESEARCH ARTICLE

Influence of surgical suture properties on the tribological interactions with artificial skin by a capstan experiment approach

Gangqiang ZHANG[1,2,3], Tianhui REN[3], Xiangqiong ZENG[1,2,*], Emile VAN DER HEIDE[2,4]

[1] *Advanced Lubricating Materials Laboratory, Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China*

[2] *Laboratory for Surface Technology and Tribology, University of Twente, Enschede 7500AE, the Netherlands*

[3] *School of Chemistry and Chemical Engineering, Key Laboratory for Thin Film and Microfabrication of the Ministry of Education, Shanghai Jiao Tong University, Shanghai 200240, China*

[4] *TU Delft, Faculty of Civil Engineering and Geosciences, Stevinweg 1, Delft 2628 CN, the Netherlands*

Received: 08 August 2016 / Revised: 26 October 2016 / Accepted: 11 December 2016

Abstract: Tribological interactions between surgical suture and human tissue play an important role in the stitching process. The purpose of the paper is to understanding the tribological behavior of surgical suture interacting with artificial skin, with respect to surgical suture material and structure, by means of a capstan experiment approach and a contact area model. The results indicated that structure and surface topography of the surgical suture had a pronounced effect on the tribological interactions. The apparent coefficient of friction of vicryl surgical suture was the smallest among the three surgical suture materials. As the sliding velocity increased, or the applied load decreased, the coefficient of friction increased. Furthermore, stick-slip phenomena were observed during the sliding procedure.

Keywords: friction; sliding; surgical suture; artificial skin; capstan experiment

1 Introduction

Surgical sutures are essential for the re-approaching of divided tissues and for the ligation of the cut ends of vessels [1], and play a significant role in wound repair by providing support to healing tissues. The physical properties, handling characteristics, and biological properties affect the basic performance of a surgical suture material [2]. Besides biocompatibility, the ideal surgical suture should (1) be easily pulled through tissues; (2) provide knotting security; (3) allow for easy handling; (4) have predictable tensile strength and performance; and (5) not shrink in tissues. With respect to the ability to be easily pulled through tissues and knotting security, the friction during surgical suture-tissue interactions plays a crucial role. Furthermore, friction generated from the surgical suture relates to the amount of "tissue-drag" and trauma [3].

Researchers have tried to improve the frictional performance of surgical sutures by various methods. Dart and Dart [2] found that the smooth surgical suture surface causes less trauma and is particularly relevant in delicate tissue like the eye. Walling et al. [4] tried to lubricate polyglactin surgical suture by using an antibiotic ointment. It was found that the lubricant could effectively decrease the coefficient of friction and improve the movement of the suture through tissue. Chen et al. [5] treated and modified braided silk surgical suture with an antibacterial agent. Both the static and dynamic friction in surgical suture-to-surgical suture contacts are affected by the antibacterial treatment, reducing friction to about 16% and 33% respectively.

There are many kinds of methods for changing the

* Corresponding author: Xiangqiong ZENG, E-mail: zengxq@sari.ac.cn

composition, structure, and surface of surgical sutures. The number of filaments and the treatment (such as a coating), has a great influence on the surgical suture properties. For instance, according to the number of filaments, surgical sutures can be divided into two kinds: monofilament and multifilament. The two kinds of surgical sutures perform differently in surgery. Monofilament surgical sutures can easily pass through living tissue due to its simple structure, resulting in little frictional resistance. Comparing with multifilament surgical sutures, monofilament surgical sutures contain little pores or interstices, which can effectively reduce the probability of harboring infectious organisms. Multifilament surgical sutures are frequently coated with the benefits of reducing frictional drag and damage to tissue, filling the interstices between fibers, and easing the repositioning of already-tied knots [6]. Three basic structures of surgical sutures, monofilament, multifilament, and coated multifilament, were studied in this paper.

The aim of this study is to understand the tribological behaviour of surgical suture, with respect to surgical suture material and surgical suture structure. Three kinds of suture materials were selected, including nylon, silk, and polygalctin. The paper is organized as follows. Firstly, we described the experiments for evaluating the friction during surgical suture sliding on an artificial skin surface. Secondly, the friction modeling of surgical sutures with different structure sliding on the artificial skin surface was presented. Thirdly, we discussed the effect of surgical suture structure and operational conditions on the tribological interactions between surgical suture and artificial skin. Finally, we drew conclusions and suggested directions for future work.

In order to study the effect of surgical suture structure and operational condition on the surgical suture friction, artificial skin, Silicone Skin L7350 [7], was used to simulate human skin. Because the skin of human and animal has a layered structure with highly changing properties through the layers [8–10], it is hard to control the friction processes at laboratory scale. And Silicone Skin L7350 was used to determinate the skin-surface friction by the Federation Internationale de Football Association (FIFA).

Surgical suture is essentially one kind of fiber used in the medical field. The classical capstan method [11–14], which has been widely used for assessing friction in fibers, was adopted in this study to investigate the frictional performance of surgical sutures. Dry and wet frictional behaviour of the surgical suture and artificial skin interface was studied in this paper, and fetal bovine serum was used in the wet frictional environment in order to simulate the stitching.

The theoretical model of the frictional behaviour of different structures of surgical suture in contact with artificial skin was set up with respect to contact area in this paper. With this modelling effort, the experimental findings could be supported. The theoretical model predicted the contact area as a function of the frictional behaviour of surgical suture.

2 Capstan measurement setup

The capstan measuring approach was used to investigate the frictional behaviour of surgical suture with artificial skin. When a textile material is pulled over a cylindrical body, a frictional force develops between material and cylinder surface due to normal force generated by tensions T_1 and T_2 exerted on the ends. The contact angle sweeps out that portion of circumference over which the contact occurs. The capstan method is designed according the Capstan Eq. (1), which has been the subject of many theoretical and experimental studies in fiber friction.

$$\mu_{\text{app}} = \ln\left(\frac{T_2}{T_1}\right)\frac{1}{\theta_{\text{wrap}}} \tag{1}$$

The apparent coefficient of friction μ_{app} of the surgical suture on the artificial skin was characterized for the suture frictional behaviour, as a function of the applied load T_1 and the tensile force T_2 in the surgical suture ends. The wrapping angle θ was π in this study.

A capstan measurement setup, as illustrated in Fig. 1, was designed by Cornelissen [15, 16], a former colleague of the group. A surgical suture is draped with an angle θ rad over a metal drum (diameter 50.0 mm). The drum is supported by mounts fitted with ball bearings at the ends of the drum on a single aluminium plate. One motor is connected to one end of the drum, which is mounted on the same base plate fitted with ball bearings. There is a motor gearhead

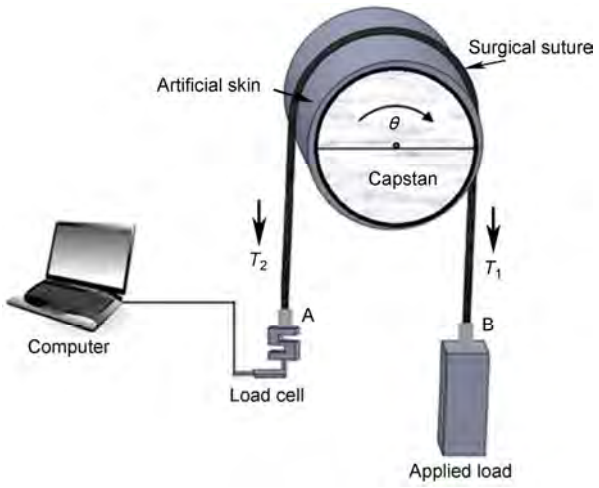


Fig. 1 Capstan experiment for friction characterization of surgical sutures.

between the drum with motor to compensate the misalignments. The force is measured by the single point load cells (range: 0–30 N) equipped with clamps. During the tests, the drum surface is covered with artificial skin by double-side tape. End A of surgical suture is fixed to the load cell; and end B is attached to an applied load.

3 Experiments

3.1 Materials

Nylon (Ningbo Medical Needle Co., Ltd.), a nylon monofilament filament; silk (Yangzhou Fuda Medical Devices CO., LTD), a braided filament; and vicryl (Ethicon, Johnson & Johnson Medical), a synthetic, braided, and absorbable surgical suture made of polyglactin 910, a copolymer of glycolide (90%) and L-lactide (10%), and coated with calcium stearate were used in this study. Silicon Skin L7350 was purchased from Maag Technic AG. Fetal bovine serum (FBS), a physiological fluids, was purchased from Sigma-Aldrich Chemie BV [17].

3.2 Experimental work

The test process was similar to that reported in the previous studies [15, 16] with respect to the frictional behaviour of fibrous tow. The artificial skin was cleaned with acetone and ethanol-impregnated textile wipe, successively before each measurement. During the experiment, the artificial skin was mounted at the cylindrical surface of the metal drum, and fixed with the double-sided foam tape. Then, a surgical suture was draped over the cleaned artificial skin, and surgical suture end A (Fig. 1) was clamped to the load cell measuring force T_2. Subsequently, an applied load T_1 was stably fixed and suspend on the end B of the surgical suture. In each test, the loading duration from the connecting of load to the start of frictional motion is controlled in about 10 s. Tests were carried out in the conditions with or without FBS. FBS was added on the surface of artificial skin at the frequency of one drop per five second. The motor, which was used to drive the drum shaft, was set to work at a rotational frequency, equivalent to a drum surface velocity of 5, 15, or 30 mm/s. The measured loads ranged from 0.1 to 2 N according to the experiments. Each test was repeated for at least three times. The apparent coefficient of friction was calculated by averaging the sampled signal at a steady state approximately 5 s after starting the motor.

3.3 Analysis of the material surface before and after the tribological test

The morphologies of the surgical sutures and the artificial skin before and after experiment were measured by Keyence VK9710 laser confocal microscopy in order to characterize the microstructure of the surgical sutures and the surface topography of the materials. In addition, the surface roughness S_a of all samples was analyzed with a laser confocal microscope VK 9710 from Keyence.

4 Contact model

4.1 Scope of the modeling approach

The model provided an intuitive description of the frictional properties of fibrous surgical suture with respect to the contact area. The friction properties of fibrous surgical suture were considered on the mesoscopic scale, with the aim to afford a relation between the structure of surgical suture and frictional behaviour. The real area of contact was calculated with a Hertzian approach [18].

When surgical suture slides over skin, the lateral

friction force acting at the skin-surgical suture interface is governed by both adhesion and ploughing component, and affected by many factors such as velocity and load.

$$F_{\mathrm{f}} = F_{\mathrm{f,adh}} + F_{\mathrm{f,pl}} \tag{2}$$

The ploughing component $F_{\mathrm{f,pl}}$ of friction in Eq. (2) plays a minor role. Neither any damage nor material transfer to the frictional artificial skin was observed. There was some wear on the surface of artificial skin shown in the Section 5.5. Therefore, for simplicity, it was suggested that the adhesion component at the contact interface between the surgical suture and the artificial skin played the dominant role. The adhesion part of the friction force $F_{\mathrm{f,adh}}$ is determined by the shear strength τ of the materials at the interface and the real contact area A_{r} between them [19]:

$$F_{\mathrm{f,adh}} = \tau A_{\mathrm{r}} \tag{3}$$

The frictional behaviour of fibrous surgical suture on artificial skin related to the real contact area A_{r} in which an interfacial shear strength τ has to be overcome during sliding. The elastic properties and the geometry of the surgical suture impacted the adhesion frictional properties on the contact area. Meanwhile, the modelling procedure with the consideration of the contact area of surgical suture on an artificial skin counterface was shown in Fig. 2.

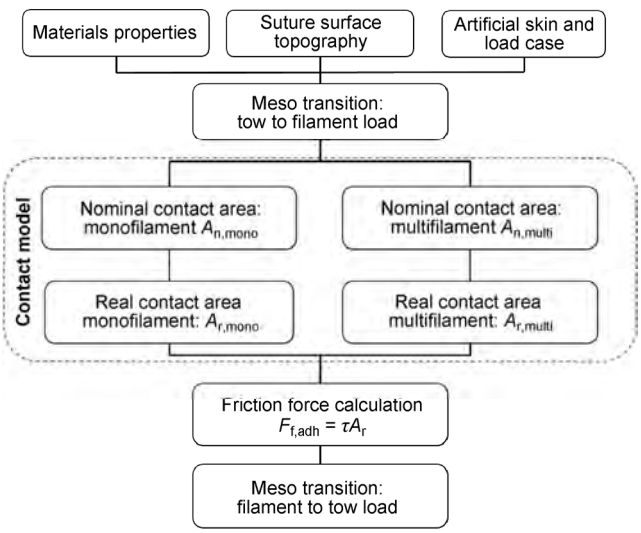


Fig. 2 Contact mechanics modelling procedure of the frictional behaviour of fibrous surgical suture contacting an artificial skin counterface.

4.2 Nominal contact area

The monofilament surgical suture was composed of a cylindrically shaped single fiber. For the multifilament surgical suture, all of the fibers were regularly twisted together to form a cylinder. The multifilament surgical sutures were assumed to be the cylindrically shaped filament without taking the structure and topography of the surgical suture into account. The nominal area of the surgical suture is formed by a line contact as shown in Fig. 3.

The nominal contact area of both monofilament and multifilament surgical suture is calculated by the Hertzian contact (Eq. 4).

$$a_{\mathrm{line}} = \left(\frac{4F_{\mathrm{N,suture}} R_{\mathrm{m}}}{\pi E^{*}} \right)^{1/2} \tag{4}$$

where a_{line} is the half-width of contact and $F_{\mathrm{N,suture}}$ represents the distributed normal load on surgical suture (Fig. 4). R_{m} is the effective radius of curvature, calculated as:

$$R_{\mathrm{m}} = \left(\frac{1}{R_{\mathrm{suture}}} + \frac{1}{R_{\mathrm{flat}}} \right)^{-1} \tag{5}$$

with the surgical suture radius of curvature in transverse direction, the counterface is represented by a flat artificial skin surface, $R_{\mathrm{flat}} = \infty$, thus $R_{\mathrm{m}} = R_{\mathrm{suture}}$.

The equivalent Yong's modulus E^{*} is defined as:

$$E^{*} = \left(\frac{1-v_1^2}{E_1} + \frac{1-v_2^2}{E_2} \right)^{-1} \tag{6}$$

with the fibrous surgical suture properties v_1, E_1, E_2, and v_2 for artificial skin material. From Eq. (4) the nominal contact area $A_{\mathrm{n}} = 2a_{\mathrm{line}}$ per unit filament length (unit: m²/m).

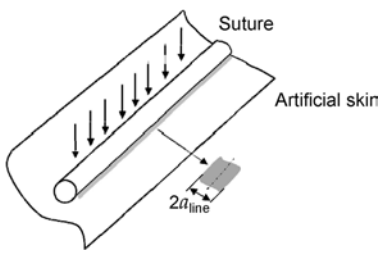


Fig. 3 Schematic drawing of Hertzian contact between a surgical suture and an artificial skin.

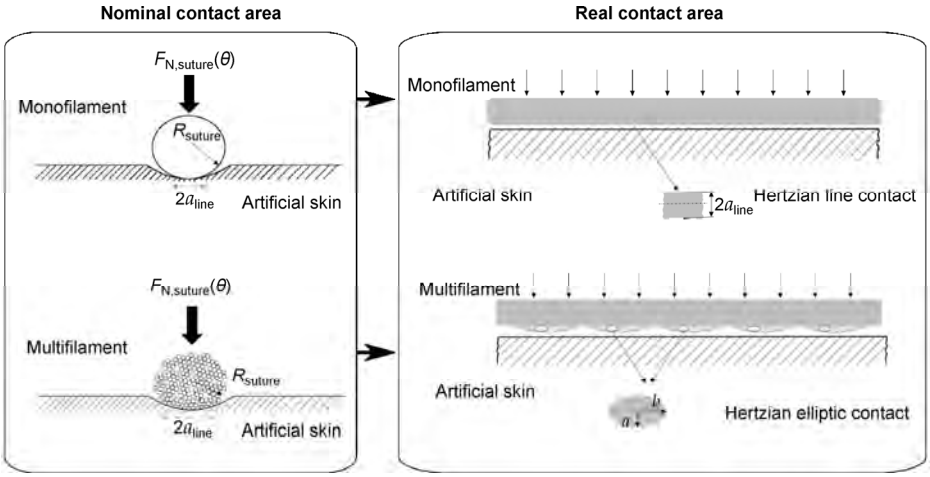


Fig. 4 Nominal and real contact area between a surgical suture and an artificial skin.

4.3 Real contact area

In general, the real contact area for the monofilament surgical suture is close to the nominal contact area as the surface of the monofilament is smoother compared with the multifilament. For multifilament surgical sutures with twisted structure and rough topography, the real contact area is different from the nominal contact area, considering the asperities on the multi-filament surgical sutures. Because the asperities on the multifilament surgical suture is formed by the certain regularly twisted fibers, the asperities could be assumed as several elliptic, as illustrated in Fig. 4. In this case, the real contact area of multifilament $A_{r,multi}$ is the sum of the asperity contact areas. For simplicity, the areas are obtained between a sphere and a cylinder with aligned principal axes by the Hertzian elliptic contact calculation [20]. The elliptic contact area equation has been derived by previous studies [15] as shown in Eq. (7) for a given distributed normal load $F_N(\theta)$, per metre surgical suture length.

$$A_{r,multi} = n_{multi} A_{multi} = n_{multi} \cdot \pi ab \tag{7}$$

where n_{multi} is the number of contacting asperities per meter filament length; a is the semi-minor axis of ellipsoid; b is the semi-major axis of ellipsoid.

5 Results and discussion

5.1 The structure and surface morphology of the materials

In this study, Keyence VK9710 laser confocal microscope was used to detect the surface topography of the surgical suture and the roughness parameters. Roughness is typically considered to be the high-frequency, short-wavelength component of a measured surface. In this research, we used arithmetic mean to describe the surface roughness. We measured the roughness in area (3D, S_a) instead of line (2D, R_a), which can give better information on the surface, for instance including the waviness profile of the filament structure. The laser confocal images of silk, vicryl, and nylon surgical sutures are shown in Fig. 5. The corresponding values of surfaces roughness (S_a) and diameters of the dry and wet surgical sutures are presented in Table 1.

From this figure, it can be seen that, besides nylon (Figs. 5(a, a_1)), both silk (Figs. 5(b, b_1)) and vicryl (Figs. 5(c, c_1)) had many filaments in their structures, and each filament in silk was thinner than that in vicryl. All filaments in vicryl were bound together due to the calcium stearate coating. Therefore, vicryl exhibited better integrity than silk in which all filaments were twisted together.

Table 1 Surface roughness and diameter of the sutures.

Items	Dry		FBS
	S_a (μm)	Diameter (μm)	Diameter (μm)
Nylon	4.25±1.01	74.78±1.23	75.59±0.89
Silk	35.13±5.31	324.81±13.84	416.62±36.34
Vicryl	14.98±2.84	74.44±2.14	87.67±1.29
Artificial skin	0.54±0.23	—	—

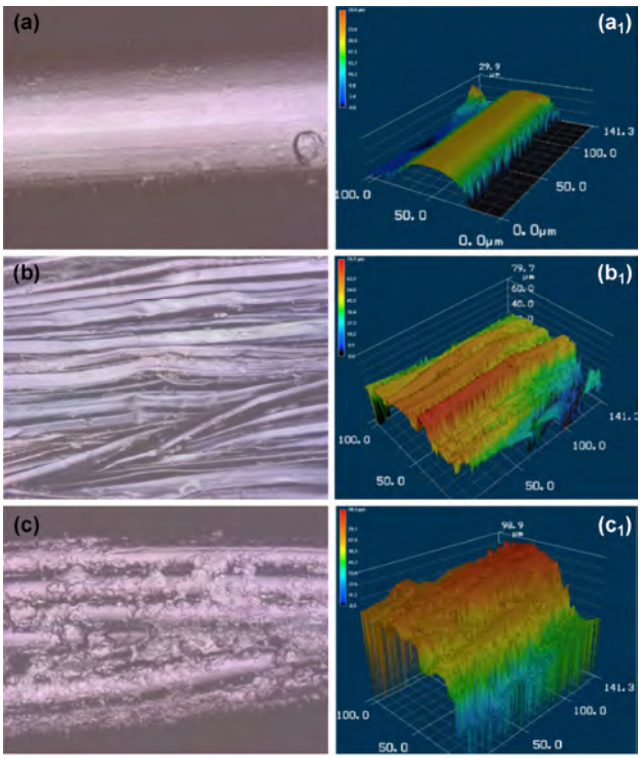


Fig. 5 Laser confocal images of the surgical suture: (a, a_1) Nylon (magnification 100×; pixel represents 0.0675 mm × 0.0675 mm); (b, b_1) silk (magnification 50×; pixel represents 0.135 mm × 0.135 mm), (c, c_1) vicryl (magnification 100×; pixel represents 0.0675 mm × 0.0675 mm).

Nylon was a kind of the monofilament and structure of Nylon is shown in Figs. 5(a, a_1). The model proposed in study is based on the assumption that the surface is smooth. The surface roughness of nylon was 4.25 μm which was the lowest in that of the three kinds of surgical sutures and that of the artificial skin was 0.54 μm as shown in Table 1. The nylon and artificial skin were assumed to be perfectly flat on the mesoscopic scale. The monofilament real contact model was applicable to this case. The real contact area for the monofilament was proportional to the square-root of the distributed normal load according to the Eq. (4).

According to the Eq. (7) mentioned in Section 4.3, the real contact area of multifilament surgical suture was determined by the number of contacting asperities per meter filament length, the semi-minor radius a, and the semi-major radius b. The surface roughness of silk was higher than that of vicryl as a whole, while much higher than that of nylon. According to the theoretical model proposed in this paper, the ratio of A_r/A_n for nylon was close to 1, while that of silk and vicryl was less than 1. Under the same experimental conditions, the contact area of nylon was larger than that of silk and vicryl. Then the friction force of nylon was supposed to be larger than that of the other surgical sutures according to the Eq. (3). Hence, friction forces of the three surgical suture are supposed to be ranked as nylon＞silk＞vicryl.

5.2 The influence of the applied load on the tribological behaviour

Figure 6 shows the value of apparent coefficient of friction at different applied loads (0.10 to 0.50 N) and the same velocity (15 mm/s) during the sliding between artificial skin and three surgical sutures. The error bars in the graphs indicate the standard deviation of each measurement. Obvious trend can be seen in this figure: the apparent coefficient of friction increased as the applied load decreased. The same trend was also observed by Viju et al. [21, 22] in case of silk-braided surgical sutures. According to the force equilibrium of the surgical suture in an infinitesimal area, the distributed normal load $F_N(\theta)$ at the contact between the fibrous surgical suture and the artificial skin counterface can be calculated as [15]:

$$F_N(\theta) = \frac{T_{\text{suture}}(\theta)}{R_{\text{drum}}} \tag{8}$$

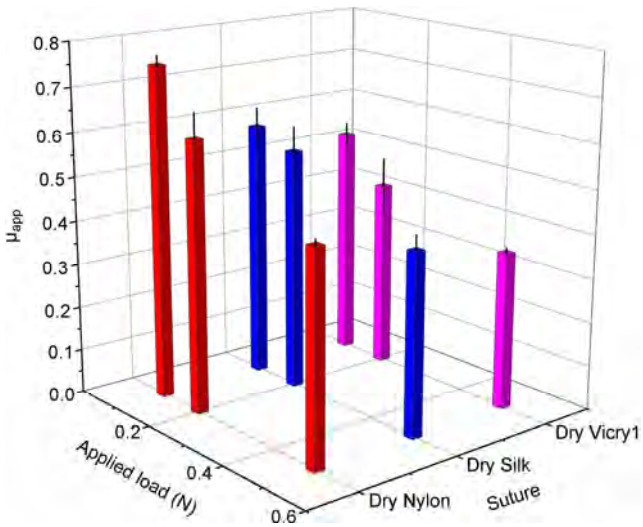


Fig. 6 Apparent coefficient of friction of (a) silk, (b) vicryl, and (c) nylon under different applied load and at 15 mm/s.

The frictional of polymer varies depending on the normal load. However, with the fibrous surgical sutures, the deformation of irregularities is elastic or it may follow an intermediate law. Fibres and polymers do not show a linear relationship between the friction force and the normal load [23]. The surgical suture is one kind of fiber which is a kind of visco-elastic materials. The normal load on the surgical suture was not directly linearly proportional to the friction force. Under higher normal load during elastic deformation, the increased friction force was not linearly proportionally with the normal load increased. During elastic deformation, the real contact area increases less proportionally with the normal load, causing the friction coefficient to decrease under higher load. In this meanwhile, the visco-elastic property of artificial skin was a significant factor to attribute to the decreasing of coefficient of friction with increasing applied load. Young's modulus E of polymer materials depends upon hydrostatic pressure in the contact zone [24]. With increasing of pressure, the modulus of elasticity for silicon is increased many times [25]. There was less deformation on the becoming harder surface of the artificial skin with increasing the Yong's modulus. The shear strength is essentially a creep carried to high deformation. So there was less shear stress when the asperities of the surgical suture slid over at higher applied load.

Another phenomenon seen from Fig. 6 is that, apparent coefficient of friction of the three adopted surgical sutures corresponding to the same normal load were ranked as nylon > silk > vicryl, which could be explained with the real contact models and the surface morphology. According to the real contact area models mentioned in the Section 4, the real contact area of nylon was the largest among the three surgical sutures. For multifilament, the uncoated filaments of silk surgical suture became flat under normal. The semi-minor radius a and the semi-major radius b of silk surgical suture were larger than that of coated vicryl in the same normal load. The contact area between silk surgical suture and artificial skin was larger than that between vicryl surgical suture and artificial skin. The increased real contact area of silk led to the results that apparent coefficient of friction of silk was higher than that of vicryl. This experimental finding verified the prediction mentioned in Section 5.1.

5.3 The influence of velocity on the tribological behaviour

Figure 7 illustrates the variation of apparent coefficient of friction with sliding velocities for silk, vicryl, and nylon subject to 0.20 N applied load. As shown in this figure, with the velocity increasing from 5 to 30 mm/s, apparent coefficient of friction of all the three surgical sutures increased, which was consistent with a similar analysis results trend in Refs. [26–28]. This phenomenon may be contributed to the viscoelastic property of the contacting materials. Textile fibers, being twisted structure and elastic performance, have time-dependent recovery behavior [29]. When load was applied, fibers increased in length, or crept and regularly flowed under the friction force. The disproportionate stress would be transferred to the weaker element of the fiber and cause them to be lengthened. The elastically deformed but unbroken bonds would return to their initial position. The regain at macroscopic dimension was time taken as hysteresis. As there was less time to regain at high velocity comparing with the case at low velocity, the apparent coefficient of friction increased with the increasing of velocity.

5.4 The influence of FBS on the tribological behaviour

Influences of FBS on the friction of three selected surgical sutures after sliding on artificial skin are shown in Fig. 8. It could be seen that, for all the three

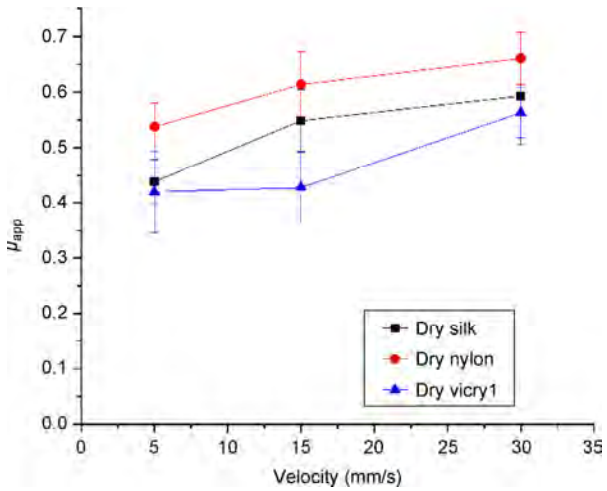


Fig. 7 Apparent coefficient of friction of silk, vicryl, and nylon at various velocity, under 0.20 N applied load.

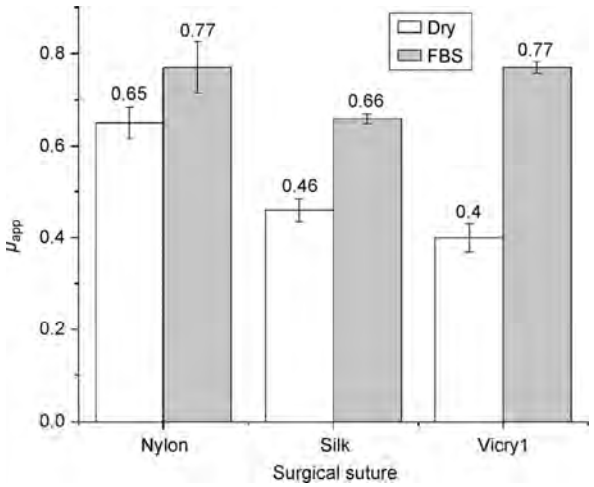


Fig. 8 Apparent coefficient of friction of silk, vicryl, and nylon at dry and the fetal bovine serum conditions under 0.10 N applied load, 15mm/s velocity.

surgical sutures, the apparent coefficient of friction in the condition with FBS was higher than that in dry condition. Reasons for this could be explained from the different contact configurations in the two conditions. There might be a thin FBS film on the contact surface. FBS is a viscous liquid collected from blood fraction. The main components in FBS are proteins and lipids, which can form an adsorbed viscous film on the surface of surgical sutures. The viscous FBS film improves the shear stress, resulting in higher friction coefficient, which was consistent with the findings of literatures [30–32].

The lager differences of apparent coefficient of friction between silk, vicryl, and nylon surgical sutures at dry condition were reduced after adopting FBS due to the absorption of moisture in the surgical suture and the formation of FBS film at the interface. The chemical properties and physical structures of surgical sutures affected their moisture absorption degrees. For instance, silk has hydrophilic group (−OH) that can form hydrogen bonds with water molecules. The twisted structure was in favor of moisture content of surgical suture.

Due to the adsorption of FBS on surgical sutures, the fibers in surgical sutures swelled. The diameters of wet silk, wet vicryl, and wet nylon infiltrated by FBS were measured by using laser confocal microscopy as shown in Table 1. The diameters of wet surgical sutures were all increased compared with the coresponding dry surgical sutures. The increased diameter of silk surgical suture, about 28.26% from 324.81 μm to 416.61 μm, was the highest in the three surgical sutures. The increased diameter ratio of vicryl was about 13.23% and that of nylon was almost no change. The results indicated that surgical suture with multifilament (silk and vicryl) had higher swelling ratio than that with monofilament as shown in Fig. 9. Moreover, coated vicryl exhibited less swelling than uncoated silk due to the better integrity of vicryl as shown in Fig. 2. The resulted increasing ratio in apparent coefficient of friction was: 88.90% for vicryl, 42.70% for silk, and 17.80% for nylon. Although vicryl had less swelling than silk, it demonstrated higher increasing in friction.

This might be due to the fact that the coating film of vicryl prevented adsorption of FBS on its surgical suture, and the coating film was swelled and became adhesive. The adhesion friction force increased due to the moist. Hence, the coating film has a significant effect on the friction when surgical suture slide through artificial skin in the condition with FBS.

5.5 Deformation of the artificial skin

Experimental observations clearly indicated that, after surgical sutures sliding over artificial skin, wear scars were generated on the artificial skin, and the wear scars were more distinct with the increasing of applied load. For low applied load, the shape and size of loaded artificial skin (Figs. 10(b, b')) are similar to that of its original form (Figs. 10(a, a')). Figures 10(c, c') showed the surface morphology and 3D morphology

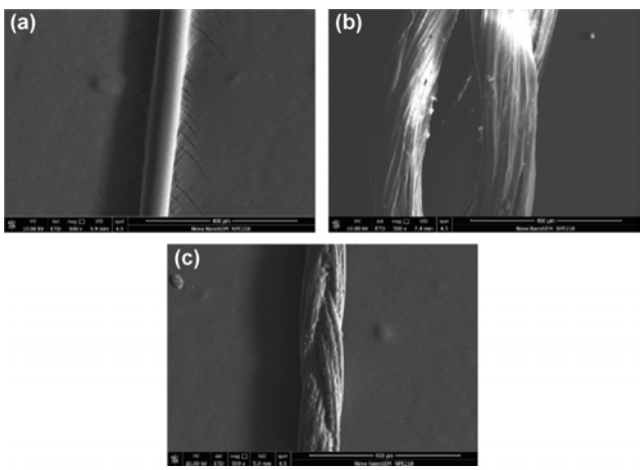


Fig. 9 SEM micrographs of surgical sutures: (a) wet nylon, (b) wet silk, and (c) wet vicryl.

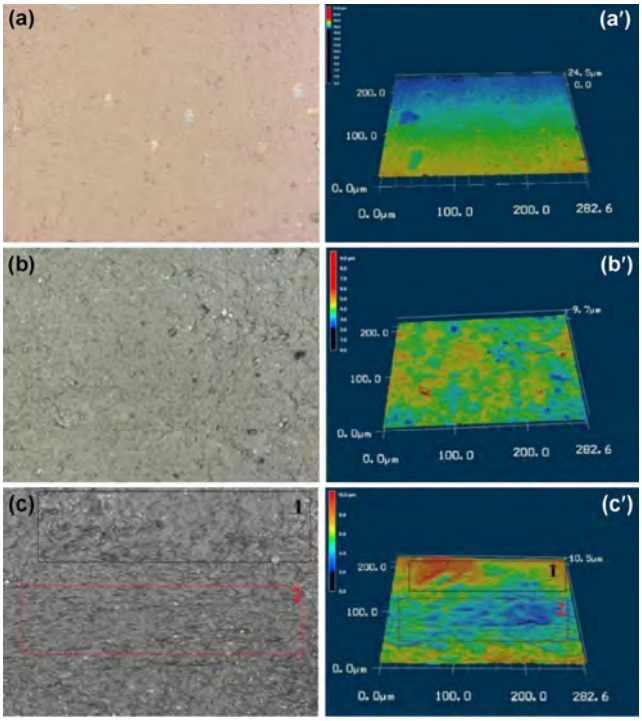


Fig. 10 Surface morphology and 3D morphology of the artificial skin; (a) the surface morphology of the artificial skin; (a') the 3D morphology of the artificial skin; (b) the surface morphology of the artificial skin tests with dry silk at 0.10 N, 5 mm/s; (b') the 3D morphology of the artificial skin tests with dry silk at 0.10 N, 5 mm/s; (c) the surface morphology of the artificial skin tests with dry silk at 2 N, 5 mm/s; (c') the 3D morphology of the artificial skin tests with dry silk at 2 N, 5 mm/s.

of the artificial skin subjected to high applied load. As shown in this figure, the surface roughness (R_a) of unworn area (marked as 1) and wore area (marked as 2) was 0.84 ± 0.06 μm and 1.19 ± 0.08 μm, respectively.

This phenomenon was explained by the influence of applied load on the tribological behavior in Section 5.2. Under high applied load, the rubbing surfaces, which were compressed repeatedly under a certain velocity and applied load, presented roughness. The generation of deformations can be explained by the deformation of the asperities on artificial skins, which could not recover completely after testing. The artificial skin experiences wear under high applied load with the increases of applied load, and the damage of asperities was more serious. Therefore, the wear scars were more distinct. The surface morphologies and 3D morphologies of artificial skin shown in Fig. 10 clearly illustrate this point.

5.6 Stick-slip patterns during sliding

The apparent coefficient of friction profile of the three surgical suture materials subject to 0.10 N applied load with the velocity of 15 mm/s in the conditions with or without FBS are shown in Fig. 11. From this figure, it could be seen that there were friction fluctuations of all the three surgical sutures in the condition without FBS, which was caused by the stick-slip in sliding. Stick-slip is a commonly observed phenomenon in fiber friction tests [12, 33].

Firstly, there was the static friction force (F_s) between surgical suture and artificial skin when there was no relative motion, and the surgical suture moved to a certain limited distant because of the deformation of the surgical suture at F_s. There was the relative motion until the tension force equals or just exceeds F_s. Secondly, the force of the surgical suture had a negative direction of the moving artificial skin because the static friction force (F_s) was less than tangential force which had the same direction with the movement of the artificial skin. At this point, the surgical suture slipped and accelerated back because of an elastic restoring force until the tension dropped to the dynamic friction force (F_d). Due to inertia, surgical suture continued to move back until it came to rest at some value of the force that was less than the dynamic friction force. Then the surgical suture started to move forward until it reached F_d. The friction force once again exceeded the elastic restoring force and the stick-slip cycle began again.

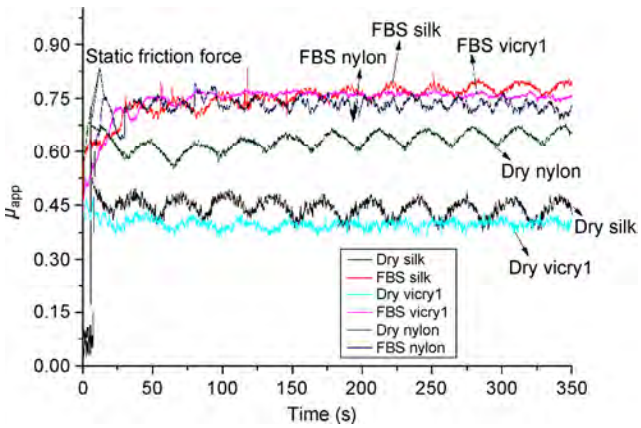


Fig. 11 The friction profile of surgical suture materials under 0.10 N applied load and 15 mm/s velocity.

Because of the difference in stick-slip pattern, the friction fluctuations of the three surgical sutures were rather different from each other (Fig. 11). For all the three surgical sutures, the amplitude of friction fluctuation of vicryl was the smallest, and the friction profile of vicryl was the smoothest in each period. There was the following possible reason for this trend. The relative velocity between surgical suture and artificial skin was different at the process of stick-slip move, and the dwell time became less at the high relative velocity. Studies indicated that contact duration or dwell time can markedly influence the level of rubber adhesion and friction [34]. When the artificial skin regularly slided, the relative velocity was changed as the tensile change.

The amplitude of friction fluctuation was significantly decreased due to the existing of FBS for the three surgical sutures, as shown in Fig. 11. This means that when surgical suture slided through artificial skin, the different between static friction force and dynamic friction force in wet condition was less than that in dry condition. When FBS was used in the experiment, FBS replaced the air and filled into the space between surgical suture and artificial skin. The three surgical sutures should overcome the same the force caused by the viscous FBS. So the coefficient of friction of the surgical sutures was similar at the present of FBS.

The less stick-slip phenomenon and amplitude of friction fluctuation occurred, the better the stitching was. It was supposed that the larger fluctuations caused greater damage to the skin. Therefore, for the three surgical sutures, the existing of FBS, which decreased the amplitude of friction fluctuation, was supposed to be benefit for the stitching.

6 Conclusion

The monofilament and multifilament contact models provided a physically understanding of the frictional behaviour between surgical suture and an artificial skin counterface. With the proposed model the experimentally observed frictional behaviour of fibrous surgical suture was explained with theoretical support.

A classical capstan method developed for the tribological study of fibers, was adopted to investigate the tribological interaction between surgical sutures and artificial skin in this study. The frictional behaviour of surgical suture was investigated by means of a quantitative comparison of three basic structure of sutures and experimental conditions. The structure and surface topography of the surgical suture was the main parameter affecting the frictional behaviour of surgical-artificial skin contact. The apparent coefficient of friction of vicryl, which was multifilament structure and coated with calcium stearate, was the smallest among the three surgical suture materials. The coated multifilament structure of the surgical suture, reduced the friction during sliding, was supposed to be benefit for the stitching. This had the certain reference significance for the future study about surgical suture development. As the surgical suture sliding velocity increased, or the applied load decreased, the apparent coefficient of friction increased. It was found that the apparent coefficient of friction increased from 0.40 to 0.77 as the fetal bovine serum added during the sliding between artificial skin and vicryl.

The stick-slip phenomena were observed in the tests with and without fetal bovine serum. The amplitude of friction fluctuation was significantly decreased due to the existing of FBS for the three surgical sutures. The existing of FBS can decrease the amplitude of friction fluctuation, but it also increased the coefficient of friction.

Acknowledgement

The authors are grateful to Marie Curie CIG (No. PCIG10-GA-2011-303922), the Shanghai Municipal "Science and Technology Innovation Action Plan" International Cooperation Project (No. 15540723600) for the financial support.

References

[1] Thacker J G, Rodeheaver G, Moore J W, Kauzlarlch J J, Kurtz L, Edgerton M T, Edlich R F. Mechanical performance of surgical sutures. *Am J Surg* **130**: 374–380 (1975)

[2] Dart A J, Dart C M. Suture material: Conventional and stimuli responsive. *Comprehensive Biomaterials* **6**: 573–587 (2011)

[3] Apt L, Henrick A. "Tissue-drag" with polyglycolic acid (Dexon) and polyglactin 910 (Vicryl) sutures in strabismus surgery. *J Pediatr Ophthalmol* **13**: 360–364 (1976)

[4] Walling H W, Christensen D R, Arpey C J, Whitaker D C. Surgical pearl: Lubrication of polyglactin suture with antibiotic ointment. *J Am Acad Dermatol* **52**(1): 136–137 (2005)

[5] Chen X, Hou D, Tang X, Wang L. Quantitative physical and handling characteristics of novel antibacterial braided silk suture materials. *J Mech Behav Biomed Mater* **50**: 160–170 (2015)

[6] El Mogahzy Y E. Development of textile fiber products for medical and protection applications. *Engineering Textiles* **2009**: 475–525 (2009)

[7] Van Der Heide E, Lossie C M, Van Bommel K J C, Reinders S A F, Lenting H B M. Experimental investigation of a polymer coating in sliding contact with skin-equivalent silicone rubber in an aqueous environment. *Tribol Trans* **53**: 842–847 (2010)

[8] Van Der Heide E, Zeng X, Masen M A. Skin tribology: Science friction? *Friction* **1**: 130–142 (2013)

[9] Morales-Hurtado M, Zeng X, Gonzalez-Rodriguez P, Ten Elshof J E, van der Heide E. A new water absorbable mechanical Epidermal skin equivalent: The combination of hydrophobic PDMS and hydrophilic PVA hydrogel. *J Mech Behav Biomed Mater* **46**: 305–317 (2015)

[10] van Kuilenburg J, Masen M A, van der Heide E.. Contact modelling of human skin: What value to use for the modulus of elasticity? *Proc IMechE, Part J: J Eng Tribol* **227**: 349–361 (2012)

[11] Robins M M, Rennell R W, Arnell R D. The friction of polyester textile fibres. *J Phys D: Appl Phys* **17**: 1349–1360 (1984)

[12] Tu C-F, For T. A study of fiber-capstan friction. 2. Stick–slip phenomena. *Tribol Int* **37**: 711–719 (2004)

[13] Gao X, Wang L, Hao X. An improved Capstan equation including power-law friction and bending rigidity for high performance yarn. *Mechanism and Machine Theory* **90**: 84–94 (2015)

[14] Tu C-F, For T. A study of fiber-capstan friction. 1. Stribeck curves. *Tribol Int* **37**: 701–710 (2004)

[15] Cornelissen B, de Rooij M B, Rietman B, Akkerman R. Frictional behaviour of high performance fibrous tows: A contact mechanics model of tow–Metal friction. *Wear* **305**: 78–88 (2013)

[16] Cornelissen B, Rietman B, Akkerman R. Frictional behaviour of high performance fibrous tows: Friction experiments. *Compos Part A: Appl Sci Manuf* **44**: 95–104 (2013)

[17] Amaral M, Abreu C, Oliveira F, Gomes J, Silva R. Tribological characterization of NCD in physiological fluids. *Diamond and Related Materials* **17**: 848–852 (2008)

[18] Johnson K L, Greenwood J A. An adhesion map for the contact of elastic spheres. *J Colloid Interface Sci* **192**: 326–333 (1997)

[19] Roselman I C, Tabor D. The friction and wear of individual carbon fibres. *J Phys D: Appl Phys* **10**: 1181–1194 (1977)

[20] Johnson K L. *Contact Mechanics*. Cambridge University Press, 1987.

[21] Viju S, Thilagavathi G. Effect of chitosan coating on the characteristics of silk-braided sutures. *Journal of Industrial Textiles* **42**: 256–268 (2012)

[22] Hendrikson W J, Zeng X, Rouwkema J, van Blitterswijk C A, van der Heide E, Moroni L. Biological and Tribological assessment of poly (ethylene oxide terephthalate)/poly (butylene terephthalate), polycaprolactone, and poly (L\DL) lactic acid plotted scaffolds for skeletal tissue regeneration. *Adv Healthc Mater* **5**(2): 232–243 (2016)

[23] Yuksekkaya M E. More about fibre friction and its measurements. *Textile Progress* **41**: 141–193 (2009)

[24] Ferry J D. Viscoelastic properties of polymers polymer solutions. *J Res Natl Bur Stand* **41**(1): 53–62 (1948)

[25] Boiko A V, Kulik V M, Seoudi B M, Chun H H, Lee I. Measurement method of complex viscoelastic material properties. *International Journal of Solids and Structures* **47**: 374–382 (2010)

[26] Roth F L, Driscoll R L, Holt W L. Friction properties of rubber. *Joumal of Research of the National Bureau of Standards* **28**: 440–461 (1942)

[27] Grosch K A. The relation between the frciton and viscoelastic properties of rubber. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **274**(1356): 21–39 (1963)

[28] Howell H G, Mieszkis K W, Tabor D. *Friction in Textiles*. London (UK): Butterworths Scientific Publications, 1959.

[29] Gupta B. Textile fiber morphology, structure and properties in relation to friction. *Friction in Textile Materials* **2008**: 3–36 (2008)

[30] Amaral M, Abreu C S, Oliveira F J, Gomes J R, Silva R F. Biotribological performance of NCD coated Si_3N_4-bioglass composites. *Diamond and Related Materials* **16**: 790–795 (2007)

[31] Briscoe B. The role of adhesion in the friction, wear and lubrication of polymers. *Adhesion* **5**: 49–80 (1981)
[32] Hansen W W, Tabor D. Role of hydrodynamic lubrication in the friction of fibers and yarns. *J Appl Phys* **27**(12): 1558–1559 (1956)
[33] Griesser H J, Chatelier R C, Martin C, Vasic Z R, Gengenbach T R, Jessup G. Elimination of stick-slip of elastomeric sutures by radiofrequency glow discharge deposited coatings. *J Biomed Mater Res* **53**(3): 235–243 (2000)
[34] Roberts A, Thomas A. The adhesion and friction of smooth rubber surfaces. *Wear* **33**: 45–64 (1975)

Gangqiang ZHANG. He received his master degree in chemistry and chemical engineering in 2013 from Qingdao University, Qingdao, China. Now, he is a Ph.D. student in the Key Laboratory for Thin Film and Microfabrication of the Ministry of Education, School of Chemistry and Chemical Engineering, Shanghai Jiao Tong University, China. Meanwhile, he is also a Ph.D. student in mechanical engineering at University of Twente, Enschede, the Netherlands. His research interests include biotribology of medical material and development of lubricants.

Xiangqiong ZENG. She received her MSc. degree in applied chemistry and Ph.D. degree in material science from Shanghai Jiao Tong University, in 2003 and 2006 respectively. After then, she worked as a staff scientist in Johnson & Johnson on Skin Care Technology for 5 years, and then as a Tenure Track Assistant Professor in University of Twente, the Netherlands, on skin and human tissue tribology for 5 years. She joined Shanghai Advanced Research Institute, Chinese Academy of Sciences from November 2015 and her current position is a professor at the Advanced Lubricating Materials Laboratory. Her research areas cover biotribology and hydration lubrication, including active control of friction and wear by surface and interface design of skin contacting materials and medical devices, by test methodology development with the design of instrument and bio-inspired human tissue model, and by additive and emulsion development for hydration lubrication.

Friction **5**(1): 99–107 (2017)
DOI 10.1007/s40544-017-0142-1
ISSN 2223-7690
CN 10-1237/TH
RESEARCH ARTICLE

Material removal mechanism of copper chemical mechanical polishing with different particle sizes based on quasi-continuum method

Aibin ZHU*, Dayong HE, Shengli HE, Wencheng LUO

Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an 710049, China

Received: 12 June 2016 / Revised: 30 August 2016 / Accepted: 16 December 2016

Abstract: In this paper, the material removal mechanism of copper chemical mechanical polishing was studied by the quasicontinuum method that integrated molecular dynamics and the finite element method. By analyzing the abrasive process of different particle sizes on single crystal copper, we investigated the internal material deformation, the formation of chips, the stress distribution, and the change of cutting force. Results showed that shear band deformation was generated along the cutting direction at approximately 45° inside the workpiece material. The deformation was accompanied by dislocations and sliding phenomena in the shear band region. Smaller abrasive particle size led to poor quality of the workpiece, while a larger particle size led to better quality. However, larger particle size resulted in greater plastic deformation and deeper residual stress inside the workpiece. Size change of abrasive particles had little effect on the tangential cutting force.

Keywords: chemical mechanical polishing; material removal mechanism; particle size; quasi-continuum; single crystal copper

1 Introduction

In IC manufacturing, copper chemical mechanical polishing (CMP) is now the most widely used planarization technology for silicon wafer processing because it achieves a super-smooth surface without injury. It is also the ideal multilayer wiring planarization method. Despite this advantage, systematic studies are still lacking, especially, in theoretical research. When comparing CMP with other technologies, CMP technology has been developed from real-world usage. The invention, development, implementation, and application of CMP have mostly occurred in industry rather than in academia. The CMP process involves tribology, chemistry, fluid mechanics, solid-state physics, and many other disciplines. Therefore, its mechanism is very complicated and increases the difficulty of research.

Precise control of the CMP process depends largely on the understanding of its material removal mechanism. When materials are removed by removing several atoms or atomic layers at a time by the CMP process, it is difficult to control and observe the process. Therefore, several difficulties exist in the experiments, calculations, and analysis of the CMP process. Steigerwald et al. provide a survey on the status of current CMP modeling [1]. They investigate different modeling assumptions and discuss the controversial treatment of physical effects. Many scholars use the molecular dynamics method to study the material removal mechanism of CMP. Yokosuka et al. developed a new accelerated quantum chemical molecular dynamics program known as "Colors" to simulate the CMP process. It is more than 5,000 times faster than the original first principles molecular dynamics program, as it is based on their original

* Corresponding author: Aibin ZHU, E-mail: abzhu@mail.xjtu.edu.cn

tight-binding theory [2]. Ye et al. simulate nanoscale polishing of a copper surface with molecular dynamics by utilizing the embedded atom method. Mechanical abrasion produces rough planarized surfaces with a large chip in front of the abrasive particle and with dislocations in the bulk of the crystal [3]. Han et al. have performed molecular dynamics analysis of chemical mechanical polishing of a silicon wafer and the simulation result shows that huge hydrostatic pressure is induced in the local area and leads to silicon atoms transforming from the classical diamond structure (α silicon) to metal structure (β silicon) [4]. Si et al. have used molecular dynamics simulations of a nanoscratching processes to investigate the removal mechanism of single crystalline silicon in the CMP process on the atomic-scale. The simulation results under a scratching depth of 1 nm showed that a thick layer of silicon material was removed by chip formation and an amorphous layer was formed on the silicon surface after nanoscratching [5]. Wu et al. investigated and calculated the wafer topography effects on the contact pressure distribution during CMP with a 3D solid-solid contact model. They proposed a formula that helps to specify the polishing parameters during CMP [6]. Zhou et al. used mixed elastohydrodynamic lubrication model to investigate the CMP process and the results showed that the proposed layer elastic theory is an optimal model to describe the polishing pad deformation behavior in CMP [7]. Chen et al. used molecular dynamics simulation to analyze the material removal mechanism of a silicon substrate under the impact of a large porous silica cluster with different pore diameters. Their findings were inspiring for other researchers who want to optimize the process parameters in order to obtain a lower surface roughness and higher material removal rate during the CMP process [8]. Lee et al. proposed a mathematical model-based evaluation method to determine the environmental burden of the CMP process. They adopt our previously reported material removal rate model for CMP and modified it to incorporate the effect of the slurry flow rate and process temperature. The results showed that slurry consumption strongly impacts the carbon dioxide equivalents of the CMP process [9]. Yang et al. proposed a chip-scale CMP model that is based on the Greenwood–Williamson theory. It combines the effects of both the pad bulk deformation and the feature size, which are mainly responsible for global pressure distribution and the feature-scale removal rate variation, respectively [10].

However, the simulated scale range of the CMP process by molecular dynamics method is limited. The maximum grain radius for simulation generally cannot exceed 10 nm, while the actual average grain radius in CMP process is 50 nm according to the work of Larsen-Basse et al. [11]. The quasicontinuum (QC) method, a multiscale simulation method that integrates the molecular dynamics method and finite element method, is potentially an effective way to solve this problem based on the work of Tadmor [12].

The QC method, proposed by Tadmor et al. [12], is one of the multiscale simulation methods. By mixing a continuum and atomistic approach, this method uses an atomic scale solution near dislocation core regions while using "representative atoms" to conduct a rough description away from the dislocation core regions. By calculating the system's energy and force in the state of a reduced degree of freedom, the mixing of atoms and continuum is achieved. In this paper, a CMP multiscale computer simulation model is used to simulate the CMP process of the microparticle contact interaction with the workpiece, which is the mechanical action process of the CMP material removal mechanism.

Figure 1 depicts the basic principle of the QC method. The black filled circles in Fig. 1(a) are the representative atoms, and Fig. 1(b) shows the finite element mesh that is composed of the corresponding representative atoms. As shown in Fig. 1(b), atoms B, C, and D are finite element nodes, and atom A is a nonrepresentative atom. The criteria for judging whether an atom is a representative atom or nonrepresentative atom is based on the formula, $\max_{a,b,k}\left|\lambda_k^a-\lambda_k^b\right|>\varepsilon$, where λ_k^a is the $k^{\text{th}}-$order eigenvalue of the right Cauchy–Green tensor $U_a=\sqrt{F_a^TF^T}$ of unit a and ε is a parameter set by practical experience. If an atom satisfies $\max_{a,b,k}\left|\lambda_k^a-\lambda_k^b\right|>\varepsilon$, then it is a representative atom; otherwise, it is a nonrepresentative atom.

The displacements of nonrepresentative atoms are

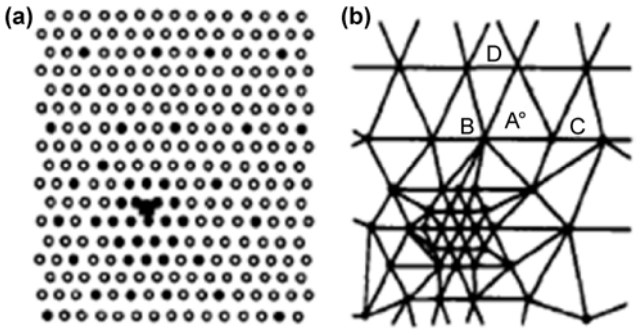


Fig. 1 Schematic of QC method.

derived from the linear interpolation for representative atoms. This purpose can be achieved using the interpolated equations of the finite element method in the QC method. In regions where a fully detailed atom description is required, we can choose all atoms as representative atoms and reduce the density of representative atoms in regions with a smaller deformation gradient. As shown in Fig. 1, we may choose all atoms in the dislocation core region as representative atoms and reduce the density of representation atoms in regions away from the dislocation core region.

Currently, the QC method has been used to study a variety of basic phenomenon in crystal deformation such as fracture, grain boundary structure and deformation, nanoindentation, and interaction of three-dimenssional dislocations. Miller et al. has considered fracture in two distinct settings. First, cracks in single crystals are investigated, and are then considered as a crack advancing towards a grain boundary in its path. In the investigation of single crystal fracture, the competition between simple cleavage and crack-tip dislocation emission is evaluated [13]. Shenoy et al. have shown how their mixed atomistic and continuum analysis is adapted to the treatment of interfacial deformation. Such calculations demand the generalization of the original quasi-continuum formulation to allow for the existence of more than one grain simultaneously [14]. Tadmor et al. investigate two different crystallographic orientations of the film and two different indenter geometries, a rectangular prism and a cylinder. They obtain both macroscopic load versus indentation depth curves and microscopic quantities, such as Peierls stress and density of geometrically necessary dislocations beneath the indenter [15].

2 Simulation model

A QC program was adopted in this paper to calculate the core, and the simulation was achieved by secondary development on that basis. The entire program includes parameters to calculate meshing, boundary conditions and outer loads, the QC computation core, and output result modules, as shown in Fig. 2.

In the CMP process, part of the abrasive particles

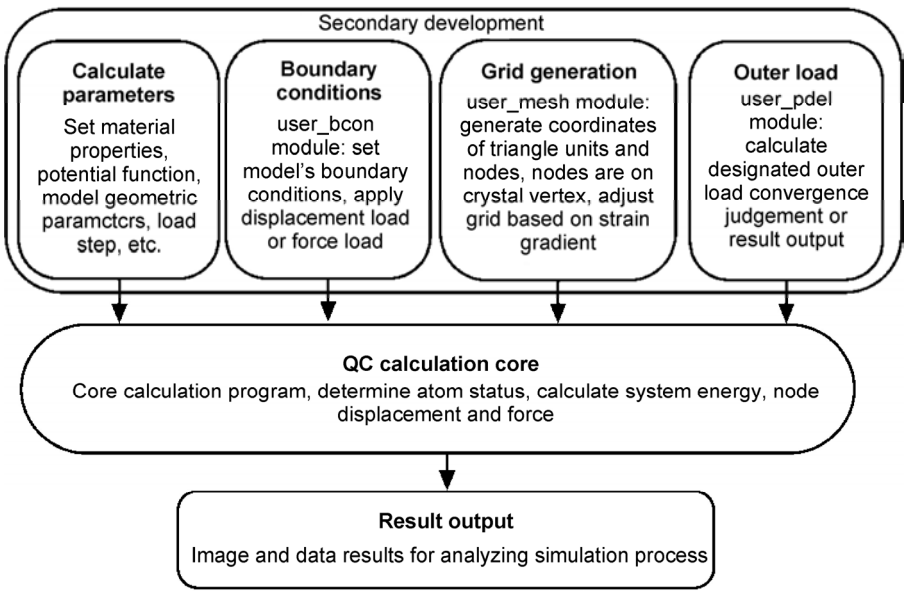


Fig. 2 QC module.

are embedded into the surface of the polishing pad. Each embedded particle is the equivalent of a fixed abrasive, plowing workpiece surface that removes workpiece material, as shown in Fig. 3.

This single polishing process was simulated in the paper, and the simulation model was shown in Fig. 4. As shown in Fig. 4, a single crystal copper workpiece was meshed in a finite element grid. The top grid was meshed denser and denser until it became an atomic region, which achieved the integration of the atomistic approach and finite element method. The workpiece dimension was 500 Å × 250 Å. Perpendicular to the paper, the width of a layer of lattice was 3.6 Å. The workpiece contained 40,000 atoms, while the number of representative atoms required for computing was less than 5,000, greatly reducing the computation amount.

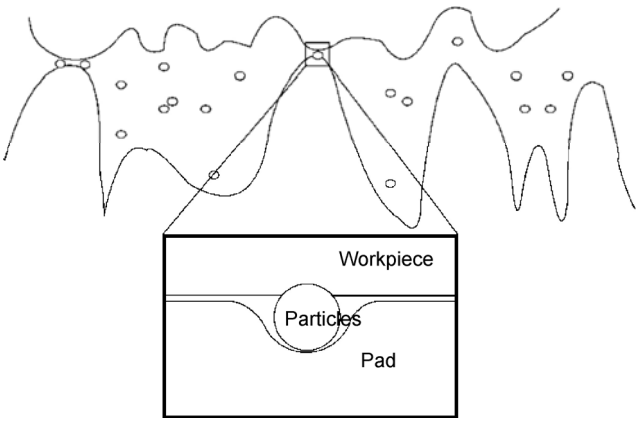


Fig. 3 CMP material removal model.

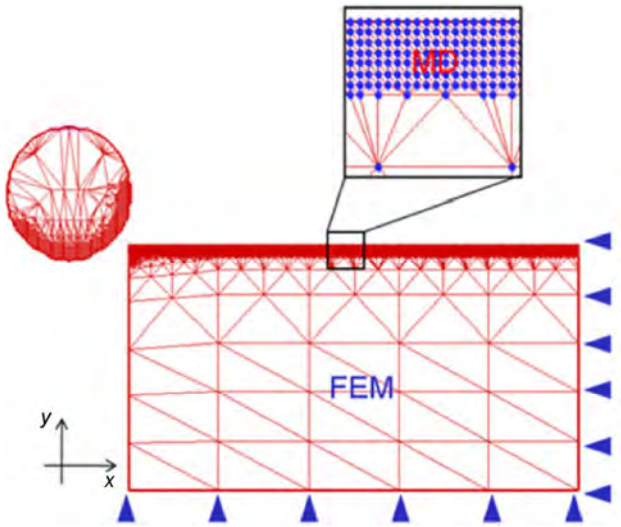


Fig. 4 Multiscale simulation model.

Ideally, in order to calculate the total energy, all atoms in the domain need to be visited by:

$$E_{\text{tot}} = \sum_{i=1}^{N} E_i(x_1, x_2, \cdots, x_N) \tag{1}$$

where E_i is the energy contribution from site x_i.

The precise form of E_i depended on the potential function used. In the regions where the displacement field was smooth, keeping track of each individual atom was unnecessary. After selecting some representative atoms, the displacement of the remaining atoms could be approximated by linear interpolation.

In the equation below,

$$u_j = \sum_{\alpha=1}^{N_{\text{rep}}} S_\alpha(x_j^0) u_\alpha \tag{2}$$

where the subscript α identified the representative atoms, S_α was an appropriate weight function, and N_{rep} was the number of representative atoms involved.

Although this step reduced the degree of freedom needed for calculation, every atom still needed to be visited. To reduce the computational complexity involved in obtaining the total energy, several simplified rules were introduced. Among these rules, the Cauchy–Born rule assumed the deformation gradient, $A = \partial x / \partial x^0$, was uniform within each element. The strain energy in the element Ω_k could be written approximately as $\varepsilon(A_k)|\Omega_k|$ in terms of the strain energy density $\varepsilon(A)$. With these approximations, the evaluation of the total energy was reduced to a summation over the finite elements:

$$E_{\text{tot}} \approx \sum_{i=1}^{N_e} \varepsilon(A_k)|\Omega_k| \tag{3}$$

where N_e is the number of elements.

Equation (3) is called the local version of QC. In the presence of defects, the deformation tends to be nonsmooth. Therefore, the approximation made in the local QC is inaccurate. A nonlocal version of the QC has been developed in which the energy is expressed as:

$$E_{\text{tot}} \approx \sum_{\alpha=1}^{N_{\text{rep}}} n_\alpha E_\alpha(u_\alpha) \tag{4}$$

where n_α is a chosen suitable weight, E_α is the energy

from each representative atom, and it is computed by visiting the representative atom's neighboring atoms whose positions are generated by the local deformation.

Practical implementation usually combines both local and nonlocal versions of the method. A criterion has been proposed to identify the local/nonlocal regions, allowing the whole procedure to be adaptively applied. Figure 5 reviews the entire flow process of the QC method.

The composition of the abrasive particle is usually Al_2O_3, SiO_2, or CeO_2 in the actual process. The stiffness of the abrasive particle is much higher than the stiffness of the workpiece. Although the deformation of the particle is extremely small in the simulation process, the particle is considered a rigid body in the paper. Abrasive particles are initially placed at the left position of the workpiece. A displacement load is applied with a load step of 0.3 Å. There are 300 steps in total. The workpiece (001) is cut along the direction of the crystal plane [100] with a cutting depth of 10 Å. Fixed constraints are applied at the bottom and the right end surfaces of the workpiece. For the direction perpendicular to the paper, a periodic boundary condition is adopted. The simulation process is quasi two-dimensional. In fact, due to the static properties

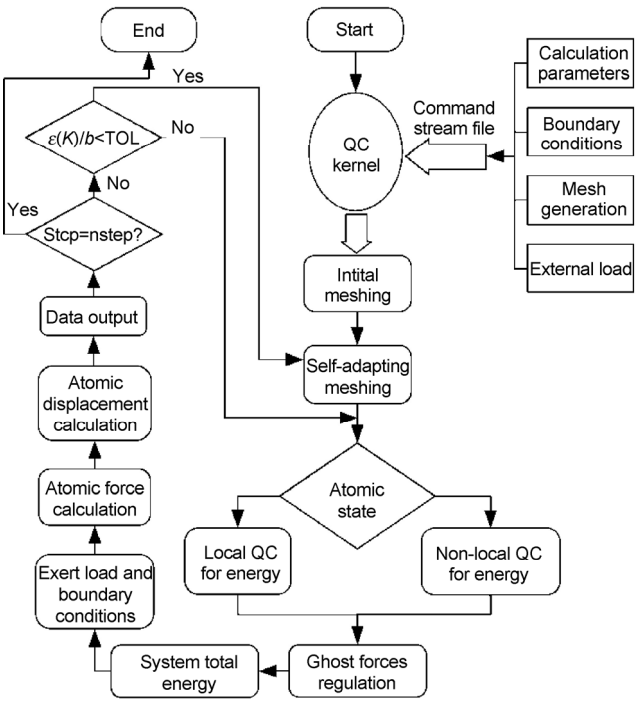


Fig. 5 Flow chart of the QC simulation.

of the QC method, energy minimization is executed every time a load step is applied. As a result, the system is in an equilibrium state at each load step. The stiffness of abrasive particles is much higher than of the workpiece. Therefore, the particles are seen as a rigid body due to little deformation in the simulation process. The radiuses of abrasive particles for the simulation are 5 nm, 10 nm, and 20 nm. In comparison with the molecular dynamics method, the particle sizes in this paper are closer to the actual situation. For the copper workpiece, the embedded atom method (EAM) potential function is adopted with a cutoff radius of 4.95 Å. The EAM potential function that describes the interaction between copper atoms can be denoted by:

$$U=\sum_i F_i(\rho_i)+\frac{1}{2}\sum_{i\neq j}\phi_{ij}(r_{ij}) \tag{5}$$

where U represents the system's total energy, $F(\rho)$ represents the system's embedded energy, ρ_i represents the sum of the electron cloud density at atom i by other extranuclear electrons, $\phi_{ij}(r_{ij})$ represents the potential energy between atom i and atom j, and r_{ij} represents the distance between atom i and atom j. $\phi_{ij}(r_{ij})$, ρ_i, and $F(\rho)$ can be denoted by Eqs. (6)–(8), respectively.

$$\Phi_{ij}(r_{ij})=\frac{A\mathrm{e}^{-\alpha(r_{ij}/r_\mathrm{e}-1)}}{1+(r_{ij}/r_\mathrm{e}-\kappa)^m}-\frac{B\mathrm{e}^{-\beta(r_{ij}/r_\mathrm{e}-1)}}{1+(r_{ij}/r_\mathrm{e}-\lambda)^n} \tag{6}$$

$$\rho_i=\sum_{j\neq i}\frac{f_\mathrm{e}\mathrm{e}^{-\beta(r_{ij}/r_\mathrm{e}-1)}}{1+(r_{ij}/r_\mathrm{e}-\lambda)^n} \tag{7}$$

$$F(\rho)=\begin{cases}\sum_{i=0}^{3}F_{ni}\left(\frac{\rho}{\rho_n}-1\right)^i, \rho<\rho_n, \rho_n=0.85\rho_\mathrm{e}\\ \sum_{i=0}^{3}F_i\left(\frac{\rho}{\rho_\mathrm{e}}-1\right)^i, \rho_n<\rho<\rho_0, \rho_0=1.15\rho_\mathrm{e}\\ F_0\left(1-\ln\left(\frac{\rho}{\rho_\mathrm{e}}\right)^\eta\right)\left(\frac{\rho}{\rho_\mathrm{e}}\right)^\eta, \rho_0<\rho\end{cases} \tag{8}$$

The simulation program runs on a Linux system, and approximately 100 hours were required for a computational process. Slightly more time was consumed when the particle radius increased. The model's parameters were shown in Table 1.

Table 1 Model parameters.

Items	Parameters
Material of work piece	Face-centered cubic single crystal copper
Crystal orientation of work piece	x[100]y[001]
Radiuses of abrasive particle	5 nm, 10 nm, 20 nm
Cutting depth	10 Å
Work piece dimension	500 Å × 250 Å
Scale of work piece	About 40,000 atoms

3 Results and discussions

3.1 Cutting chips analysis

Figures 6–8 show the atomic displacement images of the cutting process at the 50th and 200th load step with particle radiuses of 5 nm, 10 nm, and 20 nm, respectively. We can see from Fig. 6, the atomic lattices at the contact region of the workpiece and abrasive particles undergo initial deformation, and expand into a shear band along the 45° direction as the particles move forward in the plowing process. The deformation is generated mainly by compression and shear action. With the particles moving, dislocation and slipping occur in the deformation region. When the strain energy stored in the deformed crystal lattice surpasses a certain value, atoms rearrange to release the unnecessary energy, and then return to the equilibrium state. When the strain energy is insufficient to rearrange the atoms, some atomic region can move in block to present a sidestep shape on the workpiece surface.

When the particle radius is 5 nm, as shown in Fig. 6, chips can move forward as the particles move and they accumulate in front of the particles, blocking the forward way of the particles. As the particle radius increases, chips obviously no longer accumulate. That phenomenon results from the decreased angle between

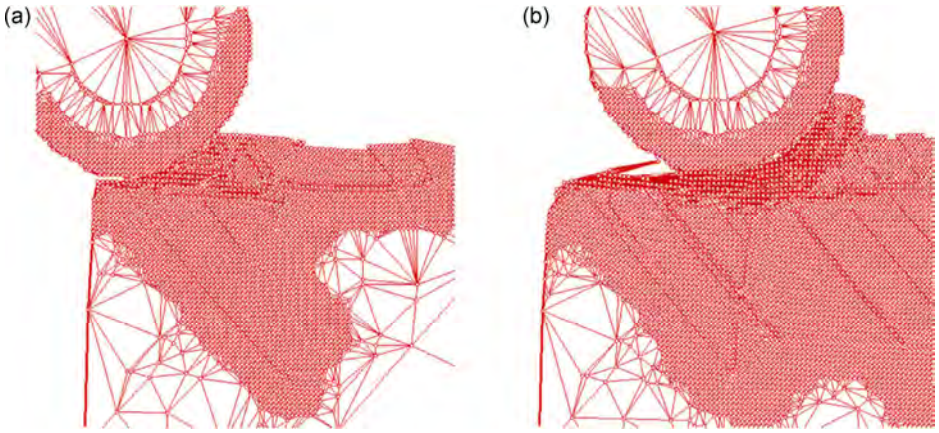


Fig. 6 Abrasive particle with radius of 5 nm at (a) 50th time step, and (b) 200th time step.

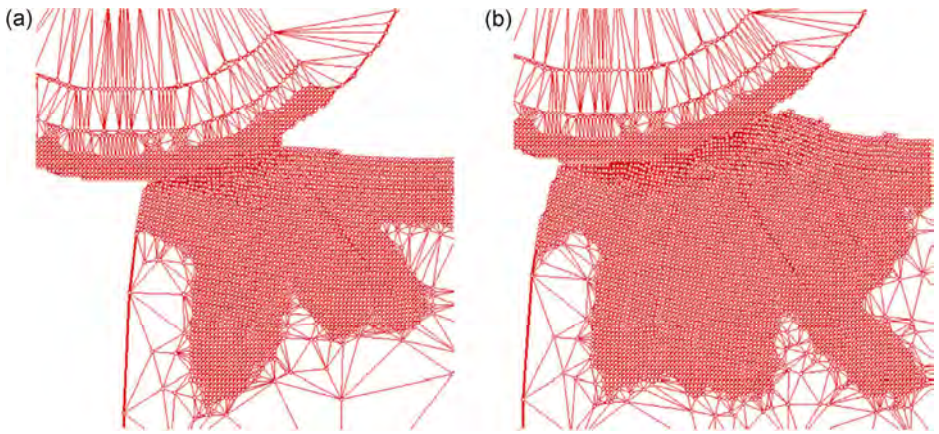


Fig. 7 Abrasive particles with radius of 10 nm at (a) 50th time step, and (b) 200th time step.

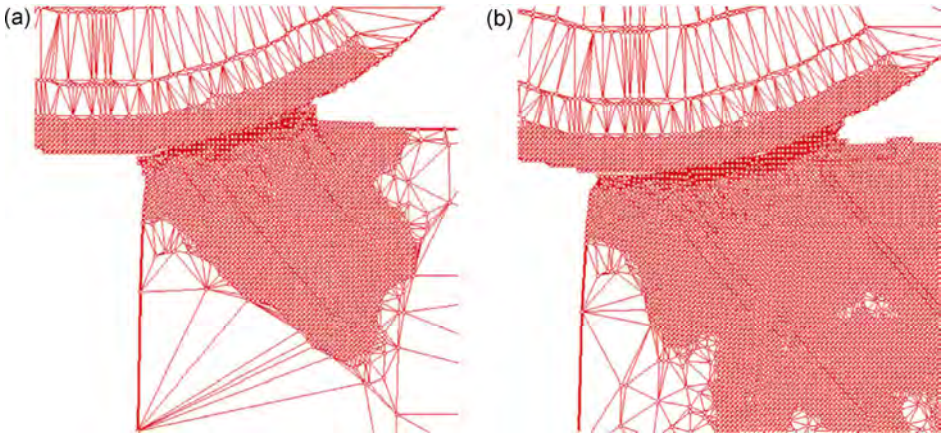


Fig. 8 Abrasive particles with radius of 20 nm at (a) 50th time step, (b) 200th time step.

the particles and the workpiece, causing the chips in front of the particles to undergo a squeezing action. As a result, workpiece material is removed by the shearing action of the particles, and the quality of the processed work piece surface is relatively poor.

As the radius increases as shown in Figs. 7 and 8, the workpiece material undergoes the shearing and squeezing actions from the particles, which leads to better surface quality and deeper deformation inside the work piece. Figure 7 and 8 show the same phenomena as described in Fig. 6.

Additionally, we can now know the advantage the QC method has in efficiency. The wafer system in this paper contains approximately 4×10^4 atoms, but the number of representative atoms needs to be calculated in an atomic region less than 5,000 for the QC method. Therefore, the calculation quantity can be greatly reduced in the QC method.

3.2 Stress analysis

Figure 9 shows the stress diagrams of the workpiece at the 70th load step with the particle radiuses of 5 nm, 10 nm, and 20 nm.

As shown in Fig. 9(b), most of the stress concentrates on the deformation region beneath the particles, and they are distributed along the shear band. The internal stress distribution of the workpiece is complex. In region A beneath the particles, the workpiece material undergoes plastic deformation due to compression action, creating a massive stress concentrate in this region. Only a small amount of stress is distributed in region B. In region C, a certain amount of stress is generated due to stretching action, and the stress increases when the particles move forward. Some stress is distributed in front of the accumulated region, shown in region D.

The phenomena in Figs. 9(a) and 9(c) are similar to those in Fig. 9(b), except as particle radius increases, compression of the particles on the workpiece material becomes stronger and the stress distributes deeper into the internal structure of the workpiece. Unlike the other two situations, with a particle radius of 5 nm, more stress occurs in front of the chip-accumulation region, as shown in Fig. 9(a). This results from the different ways chips are accumulated.

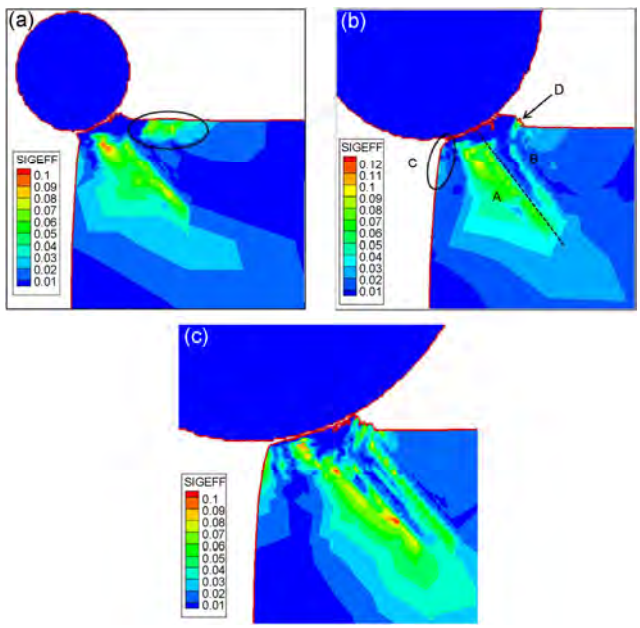


Fig. 9 Stress diagrams at the 70th load step with particle radius of (a) 5 nm, (b) 10 nm, and (c) 20 nm.

3.3 Cutting force analysis

In the QC simulation, as the particle is in uniform motion, the interaction force between the abrasive particle and workpiece is equal to the cutting force. This value can be obtained by summing the forces of all atoms that stress the particle. The cutting force is denoted as $F_{\text{cut}} = \sum_{k}^{N_{\text{cut}}} f(k)$, where N_{cut} is the number of all atoms in a particle, and $f(k)$ is the force stressed by the k^{th} atom. Figure 10 is the force–displacement plot when cutting particles of different radiuses, where the abscissa is the displacement of particles, and the ordinate is the component of the cutting force along the cutting direction.

As shown in Fig. 10, the tangential force increased rapidly with the rising cutting displacement at the initial stage of cutting. Then, the tangential force became relatively constant. The twists, turns, and sudden drops resulted from the internal dislocation nucleation of the workpiece. The average tangential cutting forces of 5 nm, 10 nm, and 20 nm particle sizes were 9.78 nN, 11.40 nN, and 8.81 nN, respectively. When the displacement was between 40–100 Å, the calculated average cutting forces of 5 nm, 10 nm, and 20 nm particle sizes were 6.75 nN, 9.36 nN, and 7.87 nN, respectively. Therefore, we could see that the change of the particle radius had little influence on the tangential cutting force.

4 Conclusions

An abrasive material removal mechanism in the copper CMP process with different particle sizes was studied in this paper. The simulation results are summarized as follows:

(1) Internal materials of the workpiece generated a shear band at approximately 45° along the cutting direction, and dislocation and sliding phenomena occurred.

(2) Chips obviously accumulated when the abrasive particles were small, and the quality of the workpiece surface was poor. Higher quality of the workpiece surface was achieved when the particles were relatively large; however, larger plastic deformation was caused inside the workpiece.

(3) The residual stresses inside the workpiece were concentrated mainly on the following regions: chip-accumulating regions along the shear band, in front of the particles, and to the left end of the processed work piece. As the particle size increased, the stress distribution deepened.

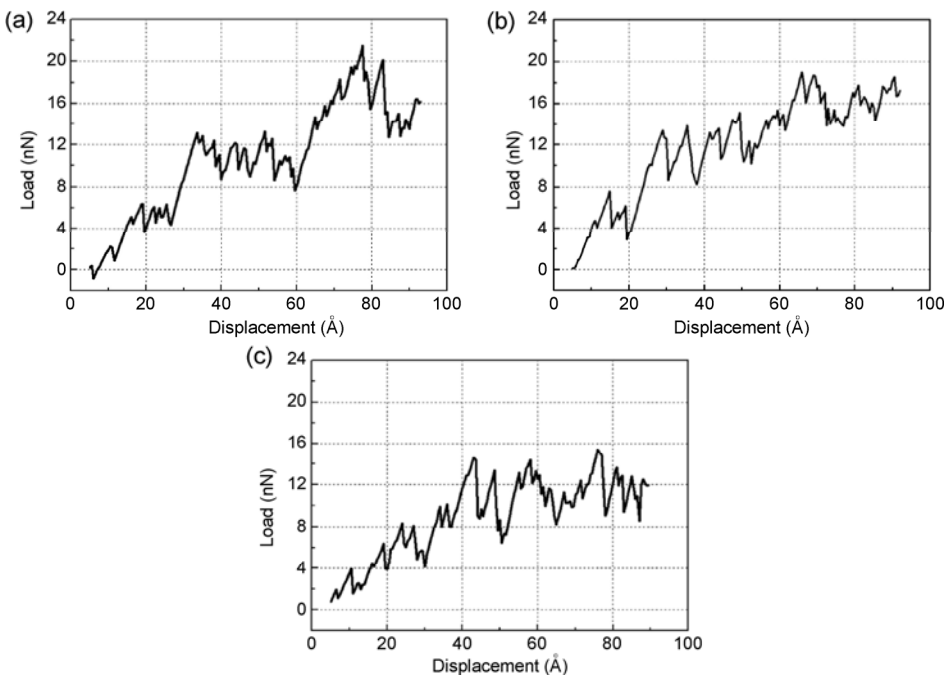


Fig. 10 Load–displacement plots with particle radius of (a) 5 nm, (b) 10 nm, and (c) 20 nm.

(4) The change of particle radius had little influence on the tangential cutting force.

Acknowledgement

The authors greatly appreciate the financial support from National Natural Science Foundation of China (No. 51175409).

References

[1] Steigerwald J, Murarka S, Gutmann R, Duquette D. Chemical processes in the chemical mechanical polishing of copper. *Mater Chem Phys* **41**(3): 217–228 (1995)

[2] Yokosuka T, Kurokawa H, Takami S, Kubo M, Miyamoto A, Imamura A. Development of new tight-binding molecular dynamics program to simulate chemical-mechanical polishing processes. *Japan J Appl Phys* **41**(4S): 2410 (2002)

[3] Ye Y, Biswas R, Bastawros A, Chandra A. Simulation of chemical mechanical planarization of copper with molecular dynamics. *Appl Phys Lett* **81**(10): 1875–1877 (2002)

[4] Han X, Hu Y, Yu S. Investigation of material removal mechanism of silicon wafer in the chemical mechanical polishing process using molecular dynamics simulation method. *Appl Phys A* **95**(3): 899–905 (2009)

[5] Si L, Guo D, Luo J, Lu X. Monoatomic layer removal mechanism in chemical mechanical polishing process: A molecular dynamics study. *J Appl Phys* **107**(6): 064310 (2010)

[6] Wu L. An analytical model of contact pressure distribution caused by 3-D wafer topography in chemical-mechanical polishing processes. *J Electrochem Soc* **159**(3): H266–H276 (2012)

[7] Zhou P, Guo D, Kang R, Jin Z. A mixed elastohydrodynamic lubrication model with layered elastic theory for simulation of chemical mechanical polishing. *Int J Adv Manuf Technol* **69**(5–8): 1009–1016 (2013)

[8] Chen R, Jiang R, Lei H, Liang M. Material removal mechanism during porous silica cluster impact on crystal silicon substrate studied by molecular dynamics simulation. *Appl Surf Sci* **264**(1): 148–156 (2013)

[9] Lee H, Dornfeld DA, Jeong H. Mathematical model-based evaluation methodology for environmental burden of chemical mechanical planarization process. *Int J Prec Eng Manuf-Green Technol* **1**(1): 11–15 (2014)

[10] Yang Z, Xu Q, Chen L. A chemical mechanical planarization model including global pressure distribution and feature size effects. *IEEE Transactions on Components Packaging & Manufacturing Technology* **6**(2): 177–184 (2016)

[11] Larsen-Basse J, Liang H. Probable role of abrasion in chemo-mechanical polishing of tungsten. *Wear* **233**: 647–654 (1999)

[12] Tadmor E B, Ortiz M, Phillips R. Quasicontinuum analysis of defects in solids. *Philosophical Magazine A* **73**(6): 1529–1563 (1996)

[13] Miller R, Tadmor E, Phillips R, Ortiz M. Quasicontinuum simulation of fracture at the atomic scale. *Modelling and Simulation in Materials Science and Engineering* **6**(5): 607–638 (1998)

[14] Shenoy V, Miller R, Tadmor E, Phillips R, Ortiz M. Quasicontinuum models of interfacial structure and deformation. *Phys Rev Lett* **80**(4): 742–745 (1998)

[15] Tadmor E, Miller R, Phillips R, Ortiz M. Nanoindentation and incipient plasticity. *J Mater Res* **14**(06): 2233–2250 (1999)

Aibin ZHU. He received his M.S. and Ph.D. degrees in mechanical engineering from Xi'an Jiaotong University, China, in 2002 and 2006 respectively. He is now an associate professor at Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University. His research areas include tribology design, modern design, and design of robot systems.

Friction **5**(1): 108–114 (2017)
DOI 10.1007/s40544-017-0148-8
ISSN 2223-7690
CN 10-1237/TH
RESEARCH ARTICLE

Tribology behavior on scratch tests: Effects of yield strength

Biao FENG[*,†]

Department of Aerospace Engineering, Iowa State University, Ames, Iowa 50011, USA

[†] *Present address: Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87544, USA*

Received: 04 October 2016 / Revised: 16 November 2016 / Accepted: 11 January 2017

Abstract: A three-dimensional (3D) scratch model is proposed to investigate the effects of yield strength of both coatings and substrates. With the help of combined Coulomb and plastic friction, the obtained results comprehensively interpret the experimental phenomena in most metals that with the growth of hardness after heat treatment the scratch friction coefficient (SFC) increases. This interpretation could not be done before. Scratch tests on the surface with or without the coating are discussed. Without the coating the SFC increases due to the decrease of the area with plastic slippage and/or the increase of friction stress during the increase of the yield strength in the material. With a softer substrate the friction stress decreases but the SFC increases, which is caused by the growth of the entire contact area and surface deformation. Conversely, with a stronger substrate the SFC decreases due to an intensified plastic slippage. The obtained results pave a new way to understanding the effects of yield strength on scratch tests, interpret experimental phenomena, and should be helpful for an optimum design in experiments.

Keywords: tribology behavior; yield strength; coating; scratch; finite element method

1 Introduction

There has been an increased interest in the application of surface coatings by changing the surface properties independently of the bulk material [1−4]. Currently scratch tests are widely employed to evaluate the tribological properties of coatings. The finite element method (FEM) becomes an important tool to describe the mechanical responses in the scratch process; it interprets the experimental phenomena; and it further develops experimental, optimum designs. Pioneering 3D FEM simulations of scratch tests, using classic elastoplasticity and Coulomb friction, have been developed for these goals [5−10].

The yield strengths of surface coatings and substrates are very important parameters, and significantly affect the materials' tribological performances. For most metals the hardness increases after heat treatment [11]. It is clear that this increase of hardness or yield strength (yield strength is approximately one third of hardness [12]) leads to a growth in the scratch friction coefficient (SFC); however, a direct explanation for this phenomenon was absent [11]. Previous FEM results based on the traditional Coulomb friction failed to explain it and even displayed an opposite trend; with a growth of yield strength the scratch friction coefficient reduces (e.g., Ref. [8, 10]). The reason was that with the growth of yield strength the deformation of surface material was suppressed, which led to a decrease of the surface deformation friction coefficient (SDFC) and caused a further drop in the SFC [13]. There was a brief attempt to explain this phenomenon in Ref. [13] which combined the Coulomb and plastic friction. One possible reason in Ref. [13] was given for a special case with soft coatings on a very hard substrate, however, it may not be true in other circumstances. As mentioned in Ref. [13], the effects of yield strength need to be reexamined and investigated in a separate

* Corresponding author: Biao FENG, E-mail: biao@lanl.gov

paper. Consequently, one of the goals of this paper is to study the effects of yield strength in the surface coating and further interpret the experimental phenomenon in detail. In addition, the yield strengths in both substrate and coating play an important role in tribological performances of the material surface. Another goal is to investigate the effects of the combined yield strengths of the substrate and coating, which will give completely distinct viewpoints from previous FEM results [5–10, 13].

2 Numerical models

A schematic diagram of a scratch system is shown in Fig. 1. The simulations in this letter follow two continuous steps: first, a vertical external force F_n is exerted along the *y*-axis to move the spherical indenter towards the coated surface; and second, a horizontal force F_t and a moment M_z are applied to move the indenter in the *x*-axis direction under a fixed F_n. Such a loading process with a constant vertical force F_n is often used in both experiments and simulations of scratch. All of the results shown in this paper will be for when the indenter slides far away from the initial indentation to avoid the effects of the initial indentation. The moment M_z along the *z*-axis is necessary to keep the indenter from rotating and was often neglected in schematic diagrams in previous literature. As stated in Ref. [7, 9], it is generally accepted that the scratch test is suitable for coatings with the thickness ranging from 0.1 to 20 μm, which covers a large number of engineering applications. The thickness of the coating and substrate is 5 μm and 35 μm respectively, and the radius of the spherical indenter is 200 μm. The width and length of the coating and substrate are not important parameters as long as they are large enough to exclude the boundary effects. Applications of the results and discussions in this paper are not limited to the current sizes in Fig. 1. If the size of such a scratch system was multiplied by *n*, then the stress distribution would be the same pattern as the current one if the applied normal force was changed from F_n to $n^2 \cdot F_n$ and the indenter moved horizontally along the *x*-axis.

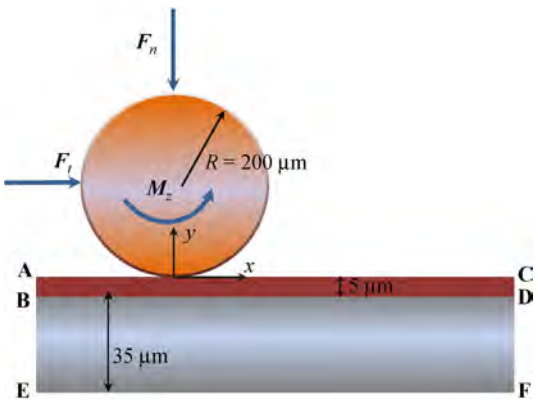


Fig. 1 Schematic representation of the scratch system including an indenter, a thin coating, and a substrate. The boundary EF is fixed during scratch tests.

The indenter is reasonably approximated to be a rigid body. The deformations of the coating and substrate are described by the position vector of the particle in the deformed state $\boldsymbol{r} = \boldsymbol{r}(\boldsymbol{r}_0, t)$, which is a function of its initial position vector $\boldsymbol{r}_0$ in the undeformed configuration and time *t*. The multiplicative decomposition of the deformation gradient $\boldsymbol{F} = \partial \boldsymbol{r} / \partial \boldsymbol{r}_0 = \boldsymbol{V}_\mathrm{e} \cdot \boldsymbol{F}_\mathrm{p}$ into symmetric elastic stretch tensor $\boldsymbol{V}_\mathrm{e}$ and plastic $\boldsymbol{F}_\mathrm{p}$ contributions is used. While we utilize the small elastic strain assumption: $\boldsymbol{\varepsilon}_\mathrm{e} = \boldsymbol{V}_\mathrm{e} - \boldsymbol{I}$ ($\boldsymbol{I}$ is the second-rank unit tensor), plastic strains and material rotations could be large. A total system of equations for the problem of linearly-elastic, perfectly-plastic flow in the coating and substrate is used as follows:

The deformation rate $\boldsymbol{d} = (\dot{\boldsymbol{F}} \cdot \boldsymbol{F}^{-1})_\mathrm{s}$ is decomposed into elastic (subscript e) and plastic (subscript p) components:

$$\boldsymbol{d} = \overset{\nabla}{\boldsymbol{\varepsilon}}_\mathrm{e} + \boldsymbol{d}_p \tag{1}$$

Hooke's law for volumetric and deviatoric parts of the Cauchy stress $\boldsymbol{T}$:

$$p = -\frac{\sigma_{xx} + \sigma_{yy} + \sigma_{zz}}{3} = -K\varepsilon_\mathrm{v}; \quad \boldsymbol{s} = 2G\, dev\boldsymbol{\varepsilon}_\mathrm{e} \tag{2}$$

Von Mises yield condition:

$$\sigma_i = \left(\frac{3}{2}\boldsymbol{s} : \boldsymbol{s}\right)^{0.5} \leqslant \sigma_y \tag{3}$$

In the elastic region:

$$\sigma_i < \sigma_y \rightarrow \boldsymbol{d}_\mathrm{p} = 0 \tag{4}$$

Plastic flow rule in the plastic region:

$$\sigma_i = \sigma_y \rightarrow \boldsymbol{d}_{\mathrm{p}} = \lambda \boldsymbol{s}\,; \quad \lambda \geqslant 0 \tag{5}$$

Equilibrium equation:

$$\nabla \cdot \boldsymbol{T} = 0 \tag{6}$$

where $\overset{\nabla}{\boldsymbol{\varepsilon}}_{\mathrm{e}}$ is the Jaumann objective time derivative of the elastic strain; p is the pressure; $\boldsymbol{s}$ is the deviator of the Cauchy stress tensor $\boldsymbol{T}$, $\boldsymbol{s} = dev\boldsymbol{T}$; ε_{v} is the elastic volumetric strain; K and G are the bulk and shear moduli respectively; σ_i is the effective stress; σ_y is material yield strength; and the parameter λ is iteratively updated by satisfaction of the von Mises yield criteria in Eq. (3). Material parameters (K, G, and σ_y) have different values for the coating and the substrate.

Similar to pioneering results [5–10, 13], the size effects in Ref. [14, 15] are not considered. The following material properties were used for the metallic coatings [16]: yield strength $\sigma_y = 234$ MPa, Young's modulus $E = 74$ GPa, $v = 0.3$, and Coulomb friction coefficient $\mu = 0.3$. In this paper, to study the effects of yield strength σ_y on coatings and substrates, we will keep the same and constant elastic properties for the coating and substrate, and vary σ_y to different values.

Using the finite element code ABAQUS 6.11, a 3D scratch process was modeled and simulated. The traditional Coulomb friction was utilized in previous FEM simulations on scratch [5–10], and admits that the relative slippage on a contact surface starts when the magnitude of the friction stress vector reaches the critical value $\mu\sigma_{\mathrm{n}}$, where σ_{n} is the normal contact pressure and μ is the traditional friction coefficient. However, for elastoplastic materials the magnitude of the friction stress is limited by shear yield strength $\tau_y = \sigma_y / \sqrt{3}$, where the von Mises yield condition is used for the materials. When friction stress reaches τ_y and is unable to increase, the material loses resistance of complete cohesion, which can initiate sliding [17–22]. This type of slippage and friction is called plastic slippage and plastic friction. During scratch the plastic friction is dominated in many cases especially for soft coating [13]. In this paper a combined Coulomb and plastic friction will be used between the indenter and coating surface, in which the sliding on the contact surface can take place when friction stress reaches a critical value $\tau_{\mathrm{crit}} = \min(\mu\sigma_{\mathrm{n}}, \tau_y)$ [17–22]. In addition, a complete cohesion is used on the contact surface between the coating and substrate.

There are two main friction coefficients. One is the scratch friction coefficient (SFC) μ_s, which is the ratio of tangential and normal resultant forces $\mu_s = F_t / F_n$ (see Fig. 1). The other is the Coulomb friction coefficient (CFC) μ, which is caused by the asperity of the contact pair and is equal to the ratio of the local friction stress and normal contact stress $\mu = \tau / \sigma_{\mathrm{n}}$ when the contact pair is under Coulomb slippage instead of plastic slippage.

3 Numerical results and discussion

First, the simplest case for scratch without a coating in Figs. 2 and 3 will be discussed. In this paper "without coating" means that in Fig. 1 the material properties of substrates are the same with the coating properties. The vertical resultant force $F_n = 0.8$ N will be used in all models. Due to the symmetry in the geometry and loading in Fig. 1, half of the structure will be used in our simulation and results. The symmetry plane is localized at the $z = 0$ plane. Figure 2 shows the distributions of the normal contact stress and the magnitude of the friction stress in one half of the contact surface from the vertical view (see the coordinate system in Fig. 2, and the symmetry plane $z = 0$ goes through the bottom boundary of each from Fig. 2(a) to 2(f)).

With a growth of the yield strength σ_y the contact size l_0 reduces due to the material hardening in Figs. 2(a)–2(c), and the SFC μ_s increases in Fig. 3, which is consistent with experimental observations [11]. The morphology of the contact surface in the elastoplastic material is the same with the surface of the spherical, rigid indenter. With the increase/decrease of the contact surface area in Fig. 2, the indentation depth d will also increase/decrease in Fig. 3. When friction stress reaches yield strength in shear $|\boldsymbol{\tau}| = \tau_y$, the slippage is plastic-slippage controlled. In Figs. 2(d), 2(e), and 2(f), the yield strength in shear τ_y of the material is 135.1, 192.8, and 250.6 MPa respectively and it is noted that in the most contact area the friction stress $|\boldsymbol{\tau}|$ in Figs. 2(d)–2(f) reaches the corresponding yield strength τ_y. Therefore plastic slippage governs at the contact surfaces in Figs. 2(d)–2(f), while the indentation depth is around or less than 1 μm which is not large. l_1 in

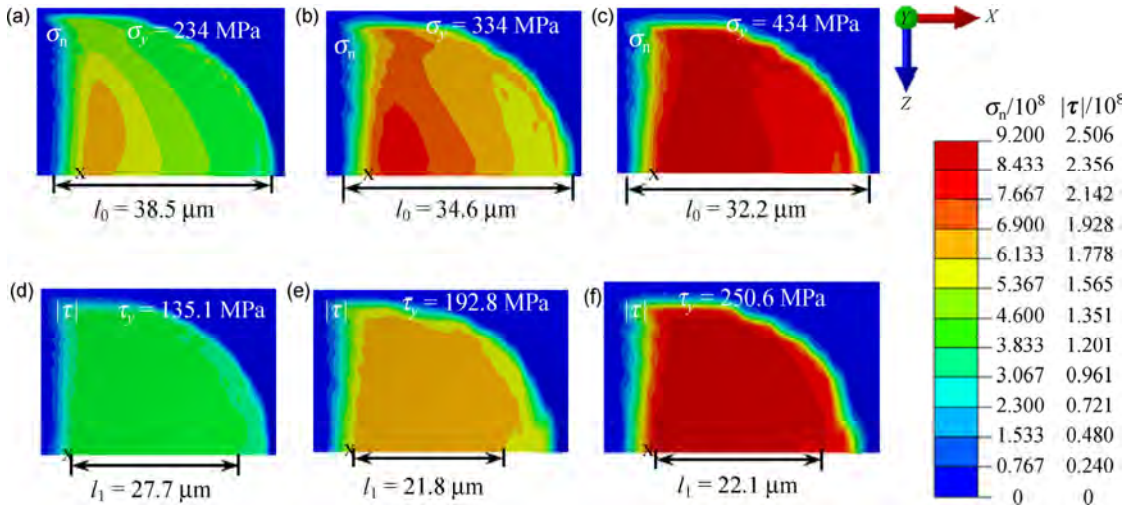


Fig. 2 Distributions of the normal contact pressure σ_n and magnitude of friction stress $|\boldsymbol{\tau}|$ on the contact surface during scratch under axial force $F_n = 0.8$ N, with a growth of yield strength σ_y or τ_y ($\tau_y = \sigma_y/\sqrt{3}$). The material properties of the coating and substrate are the same. The location of indenter tip is marked by a small x.

Figs. 2(d)–2(f) is a length parameter to represent the size of an area with plastic slippage. Material hardening occurs when an increase of the yield strength suppresses the deformation of materials and causes stress concentration. Consequently there is a much larger normal contact stress in Fig. 2(c) than in Fig. 2(a) or 2(b). The SCF $\mu_s = F_t/F_n$ in Fig. 3 is determined by the horizontal external force F_t due to a fixed normal external force F_n as 0.8 N, and the horizontal force F_t is the integral of the components of the normal stress and friction stress along the x-axis with respect to the contact area [13]. In the current cases the indentation depth is around or less than 1 μm which is quite negligible compared to the radius of the indenter $R = 200$ μm, which indicates that the contact surface is almost flat and the component of normal contact stress along the x-axis is very small. Thus the largest contribution of F_t is from the friction stress. When we neglect the bending of coating surface (i.e., the contact surface is close to the flat one), the SFC is approximately equal to $|\boldsymbol{\tau}|/\sigma_n$. The friction stress $|\boldsymbol{\tau}|$ could not increase further and stay as a constant after reaching yield strength in shear ($\tau_y = \sigma_y/\sqrt{3}$), however, the normal stress σ_n can continuously increase. It indicates that the increase of the friction slippage area may cause the reduction of the SFC and the more significant plastic slippage may cause the smaller SFC. In Ref. [13], it is mentioned that the growth of the SFC was

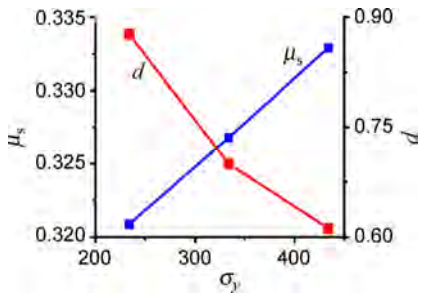


Fig. 3 Scratch friction coefficient μ_s and the indentation depth d with an increase of the yield strength σ_y under axial force $F_n = 0.8$ N in the cases when the substrate has the same material properties as the coating.

caused by a smaller contact area with friction slippage, with a growth of yield strength in the soft surface coating. One can note that this is indeed one of the reasons for SFC growth from Fig. 2(d) to Fig. 2(e), and the other reason is that the rate of growth of friction stress causing an increase in F_t surpasses the rate of the decreasing contact surface causing a drop in F_t. Obviously, the smaller area of plastic slippage in Ref. [13] with a larger yield strength in material is not the reason for the increase in the SFC from Fig. 2(e) to Fig. 2(f). One can find that the areas with plastic slippage in Figs. 2(e) and 2(f) are very comparable. Although the critical friction stress τ_y (when $\tau_y \leqslant \mu\sigma_n$) increases from 192.8 MPa to 250.6 MPa and becomes more difficult to reach on the contact surface, the stress concentration for the material with high yield

strength causes a faster growth of the friction stress, which causes that there is a similar area with plastic slippage in Figs. 2(e) and 2(f). The growth of the SFC is caused by the increasing friction stress rather than a smaller area with plastic slippage in this case.

We will discuss the case when the surface coating and substrate have different yield strengths. Figure 4 shows the distributions of the normal contact stress and the magnitude of friction stress on the coating $\sigma_y = 434$ MPa, with a substrate ($\sigma_y = 334$ MPa or 434 MPa). Here, when the coating and substrate with the same yield strength $\sigma_y = 434$ MPa in Figs. 4(b) and 4(d), it is the case for "without coating". In Fig. 4(a) presents a smaller stress concentration than in Fig. 4(b). The deformation of the soft substrate suppresses the stress concentration in the contact surface and causes a slightly larger contact surface in Fig. 4(a). With less stress concentration it is more difficult for the shear stress to reach yield stress in shear, which causes there to be a smaller area in Fig. 4(c) with plastic slippage than in Fig. 4(d). Smaller plastic slippage leads to a larger SFC for the case in Fig. 4(c), which can be seen in Fig. 5(a). One can find that in Fig. 4(d) the shear stress is slightly larger than in the one in Fig. 4(c) in a large area, however the horizontal force F_t (or SFC) is smaller in Fig. 4(d). One reason is that Fig. 4(c) has a slightly larger contact area, which may cause the integral of the friction stress to be a slightly larger than that in Fig. 4(d). The other reason is caused by a slightly larger surface deformation friction coefficient (SDFC) due to larger indentation depth in Fig. 4(a), which enlarges the SFC (see in Ref. [13]). Consequently, if the plastic slippage reduces then the SFC will increase with a soft substrate. Conversely, with a slight stronger substrate ($\sigma_y = 334$ MPa) the plastic slippage in the coating ($\sigma_y = 234$ MPa) due to stress concentration grows and leads to a reduction of the SFC, in comparison with the case without coatings, shown in Fig. 5(a). In addition, Fig. 5(a) presents that with or without a coating the growth of the yield strength in the materials causes an increase of the SFC. Figure 5(b) shows that with a growth of the yield strength of the gasket, the SFC reduces. The reason is that when yield strength increases in the substrates the stress concentration is intensified, which leads to more obvious plastic slippage in the coating surface and a smaller SFC.

4 Conclusions

In summary, a 3D scratch model is proposed to study the effects of yield strengths in both coatings and substrates by using FEM. A combination of Coulomb and plastic friction is applied on the contact surface

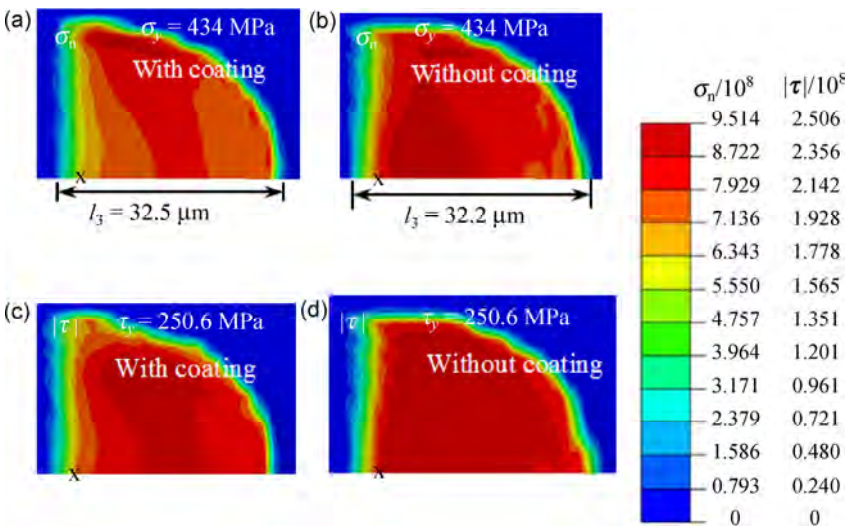


Fig. 4 Distributions of the magnitude of friction stress $|\boldsymbol{\tau}|$ and normal contact pressure σ_n on the contact surface of a coating with $\sigma_y = 434$ MPa during scratch under axial force $F_n = 0.8$ N, with a substrate ($\sigma_y = 334$ MPa in (a) and (c) and $\sigma_y = 434$ MPa in (b) and (d)). In (b) and (d), the yield strengths of coating and substrate are the same, which corresponds to the case: "without coating". The location of indenter tip is marked by a small x.

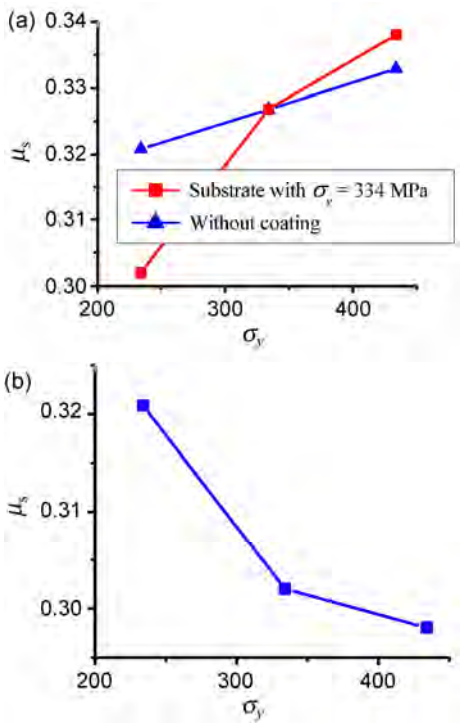


Fig. 5 (a) Variation of the SFC with respect to an increase of the yield strength in the coating with the substrate ($\sigma_y = 334$ MPa) as the red curve, and the substrate with the same yield strength of coating (i.e., "without coating") as the blue curve. (b) Variation of the SFC with a rising yield strength in the substrate for the coating with a constant yield strength $\sigma_y = 234$ MPa .

between the indenter and coating surface. The results show that with and without the coating the increase in the yield strength in the material can cause an increase of the SFC, which is consistent with experimental observations and could not be done before. Without the coating, during an increase of the yield strength in the material, the SFC grows due to a reduction of the area with slippage and/or the increase of the friction stress. When yield strength of the substrate reduces the friction stress may increase but the SFC reduces due to a reduction in contact area and a decrease of the SDFC. Conversely, when yield strength of the substrate grows the stress concentration increases on the coating surface and leads to a more obvious plastic slippage, which causes a reduction of the SFC. The deformation of elastoplasticity also has a significantly effects on the SFC, and especially it determines the SFC before the plastic slippage appears or when the plastic slippage is not dominant in the contact area. Without considering plastic slippage, the effect of the material deformation was widely studied before as in Ref. [5−10]. In the extreme case, when the indenter is deeply inserted into the coating, the effect of sever deformation of contact surface on SFC may be comparable to or even surpass the effect of plastic slippage even though plastic slippage takes place in most of contact region, which requires further study and still stays as a challenge, due to the convergence of simulation and multi-physics such as wear and fracture involved. The obtained results help in better understanding the effects of yield strength on scratch tests and to interpret experimental phenomena, and should be helpful for an optimum design in experiments.

References

[1] Bull S J, Rickerby D S, Jain A. The sliding wear of titanium nitride coatings. *Surf Coat Technol* **41**: 269–283 (1990)

[2] Chalker P R, Bull S J, Rickerby D S. A review of the methods for the evaluation of coating-substrate adhesion. *Mat Sci Eng A-Struct* **140**: 583–592 (1991)

[3] Hogmark S, Jacobson S, Larsson M. Design and evaluation of tribological coatings. *Wear* **246**: 20–33 (2000)

[4] Ivanov V V, Lebedev V A, Pinahin I A. Improving wear resistance of surface by depositing vibrational mechano-chemical MoS_2 coating. *J Frict Wear* **35**: 339–342 (2014)

[5] Bucaille J L, Gauthier C, Felder E, Schirrer R. The Influence of strain hardening of polymers on the piling-up phenomenon in scratch tests: Experiments and numerical modelling. *Wear* **260**: 803–814 (2006)

[6] Felder E, Bucaille J L, Hochstetter G. Influence of the rheology of polymers on their scratch resistance: Experimental and numerical simulation studies. *Annales De Chimie-Science Des Materiaux* **28**: 15–28 (2003)

[7] Holmberg K, Laukkanen A, Ronkainen H, Wallin K, Varjus S. A Model for Stresses, Crack generation and fracture toughness calculation in scratched TiN-coated steel surfaces. *Wear* **254**: 278–291 (2003)

[8] Jiang H, Lim G T, Reddy J N, Whitcomb J D, Sue H J. Finite element method parametric study on scratch behavior of polymers. *J Polym Sci Pt B-Polym Phys* **45**: 1435–1447 (2007)

[9] Li J, Beres W. Three-dimensional finite element modelling of the scratch test for a TiN coated titanium alloy substrate. *Wear* **260**: 1232–1242 (2006)

[10] Bellemare S C, Dao M, Suresh S. Effects of mechanical properties and surface friction on elasto-plastic sliding contact. *Mech Mater* **40**: 206–219 (2008)

[11] Zhang Z F, Zhang L C, Mai Y W. Particle effects on friction and wear of aluminium matrix composites. *J Mater.Sci* **30**: 5999–6004 (1995)

[12] Hainsworth S V, Soh W C. The effect of the substrate on the mechanical properties of TiN coatings. *Surf Coat Technol* **163**: 515–520 (2003)

[13] Feng B, Chen Z. Tribology behavior during indentation and scratch of thin films on substrates: effects of plastic friction. *AIP Adv* **5**: 057152 (2015)

[14] Chen S H, Feng B, Wei Y G, Wang T C. Prediction of the initial thickness of shear band localization based on a reduced strain gradient theory. *Int. J Solids Struct* **48**: 3099–3111 (2011)

[15] Chen S H, Feng B. Size effect in micro-scale cantilever beam bending. *Acta Mech* **219**: 291–307 (2011)

[16] Ott R D, Blue C A, Santella M L, Blau P J. The influence of a heat treatment on the tribological performance of a high wear resistant high SiAl-Si alloy weld overlay. *Wear* **251**: 868–874 (2001)

[17] Feng B, Levitas V I. Coupled phase transformations and plastic flows under torsion at high pressure in rotational diamond anvil cell: Effect of contact sliding. *J Appl Phys* **114**: 213514 (2013)

[18] Feng B, Levitas V I, Zarechnyy O M. Plastic flows and phase transformations in materials under compression in diamond anvil cell: Effect of contact sliding. *J Appl Phys* **114**: 043506 (2013)

[19] Feng B, Levitas V I, Ma Y. Strain-induced phase transformation under compression in a diamond anvil cell: Simulations of a sample and gasket. *J Appl Phys* **115**: 163509 (2014)

[20] Liu C R, Guo Y B. Finite element analysis of the effect of sequential cuts and tool-chip friction on residual stresses in a machined layer. *Int J Mech Sci* **42**: 1069–1086 (2000)

[21] Zhang H W, Zhang Z, Chen J T. The finite element simulation of the friction stir welding process. *Mat Sci Eng A-Struct* **403**: 340–348 (2005)

[22] Feng B, Levitas V I. Effects of gasket on coupled plastic flow and strain-induced plase transformations under high pressure and large torsion in a rotational diamond anvil cell. *J Appl Phy* **119**: 015902 (2016)

Biao FENG. He received his Master degree from Institute of Mechanics, Chinese Academy of Sciences, Beijing, in 2011. He earned his Ph.D. degree in engineering mechanics from Iowa State University in 2015, and afterwards did his postdoc research there for one year. Biao Feng has been conducting his research at Los Alamos National Laboratory since September 2016. His current research interest is on advanced mechanics of materials under extreme conditions.

Friction **5**(1): 115–122 (2017)
DOI 10.1007/s40544-017-0149-7
ISSN 2223-7690
CN 10-1237/TH
RESEARCH ARTICLE

Importance of surface oxide for the tribology of a Zr-based metallic glass

S. J. KANG[1,2,†], K. T. RITTGEN[1,3], S. G. KWAN[2], H. W. PARK[2], R. BENNEWITZ[1,3], A. CARON[1,2,*]

[1] *NM - Leibniz Institute for New Materials, Campus D2.2, Saarbrücken 66123, Germany*

[2] *KoreaTech, Korea University of Technology and Education, Chungnam Province 330-708, Republic of Korea*

[3] *Department of Physics, Saarland University, Saarbrücken 66123, Germany*

[†] *Present address: Department of Materials Sciences and Engineering, Zhejiang University, Hangzhou 310058, China*

Received: 31 October 2016 / Revised: 28 December 2016 / Accepted: 11 January 2017

Abstract: Thermally grown surface oxide layers dominate the single-asperity tribological behavior of a $Zr_{60}Cu_{30}Al_{10}$ glass. Increase in oxidation time leads to an increased contribution of shearing and a corresponding decreased contribution of ploughing to friction. This change in the dominating friction and wear mechanism results in an overall minor decrease of the friction coefficient of oxidized surfaces compared to the metallic glass sample with native surface oxide. Our results demonstrate the importance of creating a stable oxide layer for practical applications of metallic glasses in micro-devices involving sliding contact.

Keywords: metallic glasses; surface oxide; friction; wear; nanotribology; atomic force microscopy

1 Introduction

Metallic glasses are structurally amorphous metallic alloys whose compositions are designed to prevent crystallization upon cooling from the melt [1]. Owing to their low elastic moduli high hardness metallic glasses have been reported to provide low friction and low wear rates [2–4]. Studies comparing the macro-scale wear behavior of metallic glasses and their crystalline counter-parts or steel have reported ambiguous results. Kwon et al. [5] compared the wear behaviour of a Zr-based bulk metallic glass with hardened steel. While the friction coefficients of both materials were similar, the authors reported a significantly lower wear resistance for the Zr-based bulk metallic glass and discussed the difference of underlying wear mechanisms: brittle wear in the case of the Zr-based bulk metallic glass and adhesive wear in the case of the hardened steel. Tam and Shek reported a higher friction coefficient for metallic glasses than for their crystalline counterparts despite their higher mechanical strength and hardness [6], while Yoon et al. [7] observed an increase in hardness upon crystallization of a Ni-based metallic glass but also an increase in the friction coefficient and of the wear rate. No unique correlation between the hardness and the tribological behaviour of metallic glasses has been established so far. At the macro-scale the tribological mechanisms are complex and involve a combination of plasticity, structural relaxation and transformation [8], and material transfer [9], or mixing [10, 11] with subsequent chemical reactions and structural changes.

Meanwhile, metallic glasses have been recognized as potential materials for micro-electromechanical applications, such as micro-gears and media-storage devices [12, 13]. This potential owes to their excellent thermoplastic formability down to the nm-scale within their undercooled liquid region [14]. For such applications, the tribological performance of metallic glasses at the nanometer scale is of crucial importance. Recently, some of the authors have investigated the nanoscopic friction and wear mechanisms of a Pt-based metallic glass surface in UHV after its native oxide

* Corresponding author: A. CARON, E-mail: arnaud.caron@koreatech.ac.kr

removal in comparison to a Pt(111) surface [15]. We found that the wear-less friction coefficient was barely measurable on both surfaces with an inert diamond counter-body, while it significantly increased with a SiO_x counter-body. This observation demonstrates the effect of surface chemistry on tribological processes that may involve interfacial alloying. At the nanometer scale also, native and grown oxide surface films play an increasing role. A recent study has shown that the oxide layer on a Ni-based metallic glass enhances the wear resistance [16]. The tribological response of oxides is not unique and has been reported to differ depending on the bonding strength of oxygen with the metallic elements: a weaker bond was found to correspond to a higher friction coefficient [17].

In this work we use TEM, AFM, and XPS to characterize the structure and properties of oxide layers grown on amorphous $Zr_{60}Cu_{30}Al_{10}$ ribbons by annealing in air at 504 K for 30 min and 60 min and we report and discuss the load dependence of friction for oxide layers of different thicknesses.

2 Experimental methods and materials preparation

A $Zr_{60}Cu_{30}Al_{10}$ master alloy was prepared by melting elemental Zr, Cu, and Al with a purity > 99.9 wt.% in the corresponding weight proportion by arc-melting in Zr-gettered Ar-atmosphere. The melting process was repeated three times to ensure homogeneous mixing of the elements. Subsequently, 0.2 mm thick $Zr_{60}Cu_{30}Al_{10}$ metallic glassy ribbons were prepared by melt-spinning on a water-cooled Cu-wheel, using Ar as an injection gas. The amorphous structure of the as-spun ribbons was confirmed by X-ray diffraction (XRD) with K_a radiation and differential scanning calorimetry (DSC). Further samples from the same batch have been oxidized in air at T = 504 K for 30 min and 60 min. The temperature was selected well below the glass transition on-set temperature $T_{g,\text{ on-set}}$ = 673 K in order to avoid structural relaxation of the metallic glass [18, 19]. The surface composition and oxide layer thickness on as-spun and oxidized samples have been determined by X-ray photoemission spectroscopy (XPS) using a K-Alpha$^+$ XPS System, manufactured by ThermoFisher Scientific. In particular, the composition was determined from the peak area of the different elemental spectra and their bonding states were determined from the peak energy-shifts.

The effect of oxidation treatment in air on the surface topography was determined by tapping-mode AFM using a NanoWizard 3 Ultra AFM manufactured by JPK Instruments, Germany and a single crystalline Si cantilever (PPP-NCLR, NanoSensors, Germany). AFM scratching tests on as-spun and oxidized samples were performed at room temperature and in ambient conditions with a diamond-coated silicon cantilever (CDT-NCLR, NanoSensors, Switzerland). The normal and lateral spring constant of the cantilever, k_N and k_L, were determined from the geometrical beam model according to $k_N = \frac{Ewt^3}{4l^3}$ and $k_L = \frac{Gwt^3}{3h^2l}$, where E is Young's modulus, G is the shear modulus. The width w, the length l, and the tip height h were measured optically. The thickness of the cantilever was determined from the first bending resonance frequency of the cantilever f_0 according to $t = \frac{2\sqrt{12}\pi}{1.875^2}\sqrt{\frac{\rho}{E}}f_0 l^2$, where ρ is the mass density. The sensitivity of the photodetector S was determined by recording force-distance curves, in which the slope of the repulsive part corresponds to 1/S. The normal and lateral forces were calculated from the vertical and lateral photodiode voltages by assuming the same photodiode sensitivity in both directions according to $F_N = k_N S V_N$ and $F_N = \frac{3}{2}k_L\frac{h}{l}SV_L$, where V_N and V_L are the vertical and lateral voltages of the photodiode, respectively. For the diamond-coated cantilever used in this work we found $k_N = 48\ \text{Nm}^{-1}$ and $k_L = 6643\ \text{Nm}^{-1}$ [20, 21].

Single cycles of reciprocal scratch tests perpendicular to the cantilever axis were performed over a length l_s = 300 nm at normal force F_N = 500–7,000 nN, during which F_L was recorded. Wear scratches were imaged by tapping mode AFM to determine their width. The friction force was calculated according to $F_f = (F_{L,\text{trace}} - F_{L,\text{retrace}})/2$, where $F_{L,\text{trace/retrace}}$ are the lateral forces recorded during the forward and backward sliding motion of the cantilever during reciprocal scratching, respectively. For all measurements, the sliding velocity was v = 600 nm/s.

All results presented in this report are recorded with the same diamond-coated cantilever to directly compare the tribological performance of the different sample surfaces. All AFM-measurements were repeated twice with a different tip for each comparative series of measurements on the different samples. Given geometrical difference of the AFM-tips minor differences in the numerical values of the friction forces and the width of the scratches were observed but did not significantly affect the overall trend and the values of the friction coefficient and of the scratch hardness. This number of measurement repetition does not allow for statistical analysis. However, the authors would like to stress out the representative character of the measurements presented here. The significance of AFM-measurements relies on the idealization of a tribological contact between a smooth surface and a single asperity, i.e., the AFM-tip. For a given type of cantilevers, the geometry of the tips can be treated as equivalent, thanks to optimized production processes and controls of the manufacturers.

3 Samples characterization

The chemical composition of the oxide layers on the as-spun and the annealed samples is listed in Table 1. The bonding state of the elements was determined from the peak energy-shift. Correspondingly, Table 1 also indicates the molecular fraction of the respective oxide compounds. We find that the oxide layer mostly consists of a ZrO_2 and Cu_2O. After oxidation treatment, an increase in oxygen concentration was measured. This increase in oxygen concentration is attributed to a departure from stochiometry for ZrO_2 and Cu_2O. Both these oxides have a wide concentration range for oxygen. Increasing the oxidation time resulted in an increase in the Cu_2O molar fraction. Recently, Louzguine-Luzgin et al. have investigated the structure and growth rate of the native oxide on a $Cu_{47}Zr_{45}Al_8$ bulk metallic glass [22]. They found that after several days of exposition to ambient air the oxide layer mostly consisted of Zr_2O and Cu_2O, and a minor amount of Al_2O_3. The chemical composition of the metallic glass investigated in Ref. [22] is comparable to the one of the metallic glass investigated in this work and we expectedly find our results in good agreement with the results by Louzguine-Luzgin et al.

Figure 1 shows high-resolution TEM images on the cross section of a sample oxidized for 90 min. Both high resolution TEM images show that the structure of the underlying metallic alloy was amorphous after the oxidation treatment. However, crystallized regions

Table 1 Atomic elemental concentration of the oxide layer as a function of annealing time in air at 504 K. Also indicated are the corresponding molar fraction of oxide compounds.

	O	Al	Zr	Cu	ZrO_2	Cu_2O	Al_2O_3
As-spun	60.35	0.86	31.53	7.25	88.60	10.19	1.21
30min	75.56	1.06	18.87	4.50	87.15	10.40	2.45
60min	72.57	0.93	18.19	8.30	79.76	18.19	2.03

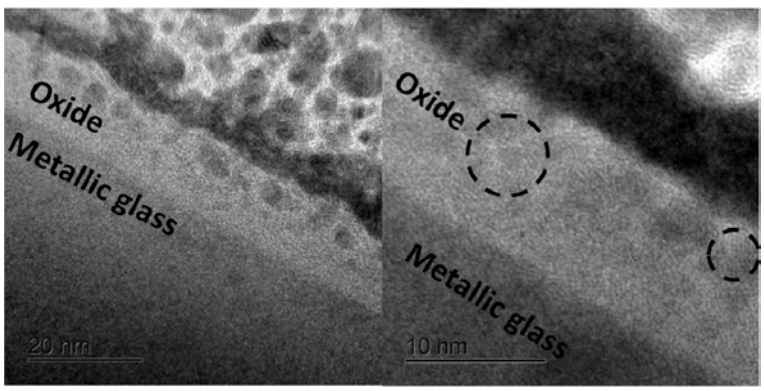


Fig. 1 Cross-sectional HRTEM images of a $Zr_{60}Cu_{30}Al_{10}$ metallic glass sample annealed for 90 min in air at 504 K. The doted circles inserted in the higher magnification micrograph (right) indicate crystalline domains.

of a few nm in diameter are observed in the oxide layer; for this sample the thickness of the oxide layer was found to be $t_{\text{oxide, 90 min}}$ = 10 nm. In Ref. [22], Louzguine-Luzgin et al. reported on the occurrence of Cu_2O nanoparticles of ~10 nm in size within a Zr-rich amorphous oxide matrix. There, the authors identified the nanocrystalline particles based on the spacing between the lattice fringes. For selected particles, they measured 213 pm, corresponding to the interplanar distance of {200}-planes in Cu_2O. In this work though, the contrast of the TEM image is to shallow to measure the spacing of the lattice fringe with accuracy. This is partly due to the fact that the TEM sample was too thick. Thus, the lattice fringes of the crystallites are overlaid with the amorphous structure of the matrix.

Figure 2 shows topography images recorded by tapping-mode AFM. The RMS roughness was between 50 pm and 80 pm. The results confirm the impressive flatness of melt-spun metallic glass ribbons, whose roughness is almost at the limit of the sensitivity of AFM experiments. It is found that the oxidation treatments did not affect the surface roughness of the samples.

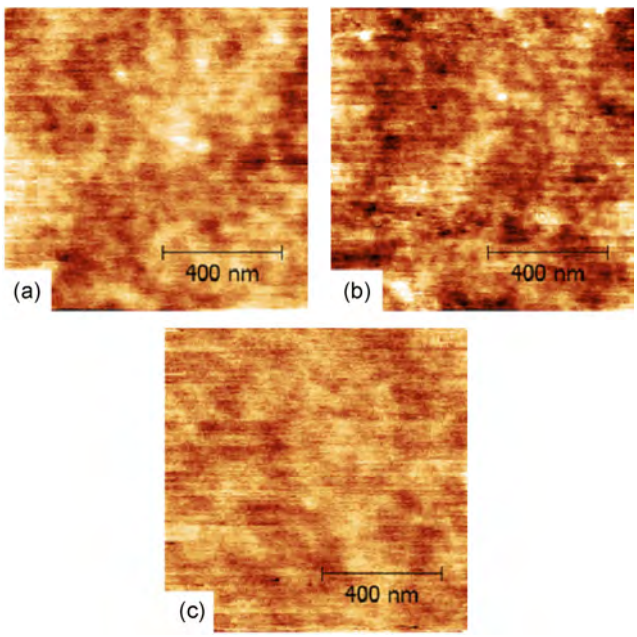


Fig. 2 Topography images recorded by tapping mode AFM on $Zr_{60}Cu_{30}Al_{10}$ metallic glass surface (a) in the as-spun state and after controlled oxidation treatment at 504 K for (b) 30 min and (c) 60 min. In all three images the *z*-scale is 500 pm. The root mean square roughness values corresponding to the topography images are (a) 71 pm, (b) 79 pm, and (c) 53 pm.

4 Results

Figure 3 shows topography images of the as-spun $Zr_{60}Cu_{30}Al_{10}$ sample after scratch tests with normal forces ranging from F_N = 0.9 μN to F_N = 4.7 μN. For normal force F_N < 2 μN, no remaining scratches could be distinguished. In this work, the projected area of contact A_c was determined from the width W_s of the remaining scratch according to $A_c = \frac{\pi W_s^2}{4}$. The scratch hardness H was determined as the inverse slope of the contact area plotted versus the normal load, i.e., $\frac{1}{H} = \frac{dA_c}{dF_N}$ or $H = \frac{dF_N}{dA_c}$. According to Bowden and Tabor [23], two mechanisms contribute to friction, the shearing of asperity junctions and ploughing. The friction force can be expressed as $F_f = \tau A_c + \mu_p F_N$, where τ is the shear strength and μ_p is the friction coefficient associated with ploughing. To decouple the shearing term from the ploughing term, the friction force has been expressed as $F_f = \tau A_c + HA_f = (\mu_s + \mu_p)F_N$, where A_f is the projected area between the submerged part of the indenter and the scratch front [23] and μ_s is the friction coefficient associated to

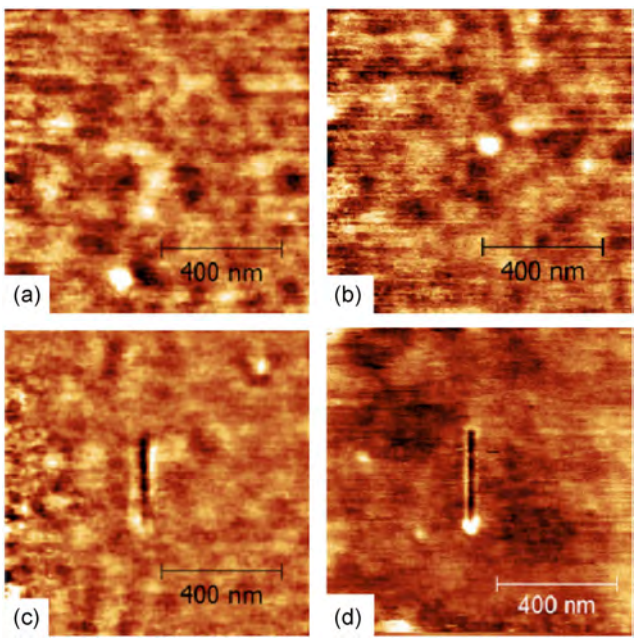


Fig. 3 Topography after scratch tests over a length l_s = 300 nm at the normal forces F_N: (a) 0.9 μN, (b) 1.9 μN, (c) 2.8 μN, and (d) 3.8 μN.

shearing. While A_c and H can be readily determined from the experimental results, the evaluation of τ and A_f is not straightforward. In principle, A_c and A_f are interdependent on W_s. For a spherical indenter A_f is approximately given by $A_f = \frac{1}{12}\frac{W_s^3}{R}$, where R is the radius of the spherical indenter.

In the wear-less regime though, the ploughing contribution can be excluded and the shearing friction coefficient μ_s can be determined from the linear slope of the friction force plotted against the normal load. In the wear regime, the linear slope of the friction force corresponds to the superposed effects of shearing and ploughing, in this case we can write $\frac{dF_f}{dF_N} = \mu_s + \mu_p = \mu$, where μ is the total friction coefficient. Figure 4(a) shows the load dependence of the friction force for the three samples investigated. For all samples remaining scratches, i.e., wear, were observed for $F_N \geqslant 2.4\,\mu N$. We observe for all three samples that the slope $\frac{dF_f}{dF_N}$ becomes steeper in the wear regime. This effect is more pronounced for the as-spun sample and less pronounced with increasing oxidation time. The friction force measured on the as-spun sample was slightly higher than on the oxidized samples over the whole range of normal forces, while it was very similar on both oxidized surfaces. The respective friction coefficient for shearing and ploughing, μ_s and μ_p, were determined as described above and are shown in Fig. 4(b). The values of the shearing friction coefficient μ_s increase from 0.177 for the as-spun sample to 0.315 and 0.309 for the samples oxidized for 30 min and 60 min, respectively. The values of the ploughing friction coefficient μ_p decrease from 0.246 for the as-spun sample to 0.102 and 0.059 for the samples oxidized for 30 min and 60 min, respectively. In the wear regime, the total friction coefficient thus changes from 0.423 for the as-spun sample to 0.417 and 0.368 for the samples oxidized for 30 min and 60 min, respectively. The value for H_s = 6.67 GPa for the as-spun sample (see Fig. 4(d)) is 1.7 times higher than those of the values for the oxidized samples of 3.95 GPa and 4.12 GPa. A clear tendency can be recognized: the surface of the investigated metallic glass mechanically softens upon oxidation. With

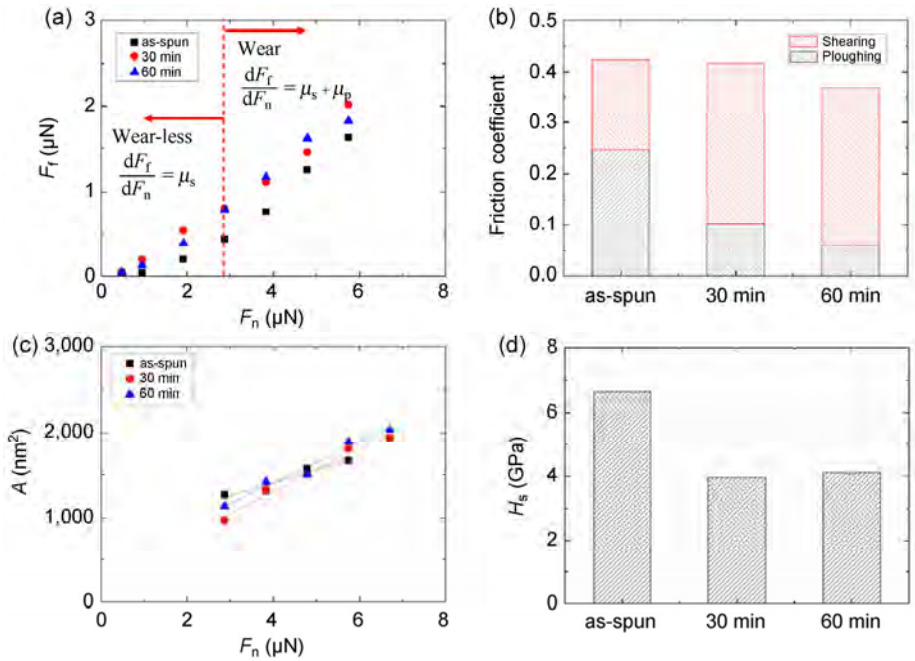


Fig. 4 (a) Friction force as a function of normal force for the as-spun and metallic glass surfaces oxidized for 30 min and 60 min. The division in wear-less and wear regime is based on the AFM results; (b) comparison of friction coefficients; (c) projected contact area as a function of the normal force; and (d) comparison of the scratch hardness.

increasing oxidation time, the contribution of ploughing to friction decreases. This observation can be explained by considering that in the case of ploughing friction forces arise from the resistance of the surface material to displacement ahead and aside of the indenter. As the hardness of a material describes its resistance to plastic flow, the effect of ploughing on friction becomes smaller for a softer surface. Moreover, the increased contribution of shearing to friction with oxidation time indicates an increased adhesion and possible material transfer.

5 Discussion

The XPS and TEM results show that thickness of the oxide layer on $Zr_{60}Cu_{30}Al_{10}$ increased up to 10 nm upon annealing in air at 504 K for 90 min. For all samples, the surface oxide was found to mostly consist in a mixture of Zr and Cu oxides. From the topography images shown in Fig. 2 with roughness lower than 1 nm we conclude that the effect of the oxidation treatment on the geometrical surface parameters can be neglected. While the effect of the surface oxidation treatment on the friction coefficient was small, our results demonstrate that the presence of a thermally grown surface oxide layer changes the governing mechanism for friction and wear. After growth of a surface oxide the contribution of ploughing to the friction coefficient decreased from 58% for the as-spun sample to 25% and 16% for the samples oxidized for 30 min and 60 min, respectively. In the later cases wear was dominated by shearing. It is interesting to note that the change in the governing mechanism for friction and wear was accompanied by a surface softening as illustrated in Fig. 5(d) by the decrease in H_s. Based on these observations we attribute the ploughing contribution to friction and wear to the plastic deformation of the metallic glass below the native or thermally grown surface oxide layer. On the other hand, shearing is attributed to adhesion of the surface oxide to the indenter. This is in line with the observation that the pile-ups on the sides and the ends of the scratches shown in Figs. 3(c) and 3(d) are much less prominent than those observed in our previous study on metallic glass surfaces in UHV conditions after oxide removal [16]. In a recent work, Louzguine-Luzgin et al. characterized the nano-scale wear of three different metallic glasses with a native surface oxide layer [24]. The authors estimated the wear rate from the scratched volume normalized by the length of the scratch performed with a sharper single crystalline diamond tip than the polycrystalline diamond coated tip used in this work. Owing to the sharpness of the tip used in Ref. [24] wear set at lower load values than in this work and the ploughing contribution is expected to be higher. Nevertheless, it is interesting to note that the authors observed a change of wear regime when the scratch depth became larger than the thickness of the native surface oxide layer. The specific wear rate became much larger when scratching deeper than the oxide layer. Though not explicitly mentioned in their work, we suggest that this corresponds to a transition from adhesive to abrasive wear. Classically, the unit less abrasive wear constant $k_{abr} = 10^{-3} - 10^{-1}$ and for adhesive wear $k_{adh} = 10^{-6} - 10^{-4}$ (see Ref. [25] for more details).

Furthermore, our experimental values for μ are lower than those found in macro-scale experiments, where the friction coefficients in the range of μ = 0.3–0.7 have been reported. The discrepancy can be explained by the absence of major plastic deformation in friction of oxide layers, which is however inherent to macro-scale friction experiments.

The tribology of ZrO_2 has been a topic of intensive research (see for example Refs. [26–28]). Recently, the low friction and wear of ZrO_2 thin layers on steel have been connected to the stress-induced martensitic transformation from tetragonal to monoclinic zirconia [26]. In our experiments, the surface oxide layer was found to be partially crystalline. The structure of the nanocrystals was not identified so that a pressure or shear induced martensitic transformation of oxide nano-grains can only be suggested as a hypothesis. Our results on the favorable tribological response of mixed Zr and Cu oxides may shed light on previous results obtained on CuO doped Y-stabilized ZrO_2 by pin-on-disc and AFM experiments, where an interfacial layer with low shear strength generated in the early stage of sliding friction tests has been invoked to explain reduced wear and friction [28].

6 Conclusions

The contribution of surface oxide layers to the single-asperity tribological response of a $Zr_{60}Cu_{30}Al_{10}$ glass increases significantly upon annealing. The thermal growth of the oxide layer leads to a decrease of scratch hardness. Shearing mediated by adhesion becomes the governing mechanism of friction and wear while ploughing related to plastic deformation of the underlying metallic glass becomes less important. The results clearly demonstrate the dominant role of the oxide layer in frictional microscopic contacts with metallic glasses and the importance of creating a stable oxide layer for practical applications in micro-devices involving sliding contact.

Acknowledgement

S. J. K., A. C. and, R. B. are grateful to Prof. E. Arzt for his continuous interest and support of this work. A. C. and S. J. K. were supported by the German Science Foundation (DFG). The authors thank Dr. M. Koch (INM) for his support by providing high-resolution TEM results.

References

[1] Inoue A. Stabilization of metallic supercooled liquid and bulk amorphous alloys. *Acta Mater* **48**(1): 279–306 (2000)

[2] Imura T, Hasegawa K, Moori M, Nishiwaki T, Takagi M, Kawamura Y. Cyclic deformation and tribological behavior of an amorphous iron-based alloy film. *Mater Sci Eng A* **133**(1): 332–336 (1991)

[3] Morris D G. The properties of dynamically compacted Metglas® 2826. *J Mater Sci* **17**(6): 1789–1794 (1982)

[4] Lee D E, Evett J E. Sliding friction and structural relaxation of metallic glasses. *Acta Metall* **32**(7): 1035–1043 (1984)

[5] Kwon D H, Lee K M, Park E S, Kim H J, Bae J C, Huh M Y. Wear behaviors of bulk metallic glass alloy and hardened steel having the same hardness value. *J Alloys Comp* **536S**: S99–S102 (2012)

[6] Tam C Y, Shek C H. Abrasive wear of $Cu_{60}Zr_{30}Ti_{10}$ bulk metallic glass. *Mater Sci Eng A* **384**(1−2): 138–142 (2004)

[7] Yoon S, Lee C, Choi H. Evaluation of the effects of the crystallinity of kinetically sprayed Ni-Ti-Zr-Si-Sn bulk metallic glass on the scratch response. *Mat Sci Eng A* **449–551**: 285–289 (2007)

[8] Boswell PG. The wear resistance of liquid quenched metallic glass. *J Mater Sci* **14**(6): 1505–1507 (1979)

[9] Kishore U S, Chandran N, Chattopadhyay K. On the wear mechanism of iron and nickel based transition metal-metalloid metallic glasses. *Acta Metall* **35**(7): 1463–1473 (1987)

[10] Fu X-Y, Kasai T, Falk M L, Rigney D A. Sliding behavior of metallic glass: Part I. Experimental investigations. *Wear* **250**(1–2): 409–419 (2001)

[11] Fu X-Y, Kasai T, Falk M L, Rigney D A. Sliding behaior of metallic glass: Part II. Computer simulations. *Wear* **250**(1–2): 420–430 (2001)

[12] Ishida M, Takeda H, Watanabe D, Amiya K, Nishiyama N, Kita K, Saotome Y, Inoue A. Fillability and imprintability of high-strength Ni-based bulk metallic glass prepared by the precision die-casting technique. *Mater Trans* **45**(4): 1239–1244 (2004)

[13] Sharma P, Zhang W, Amiya K, Kimura H, Inoue A. Nanoscale patterning of Zr-Al-Cu-Ni metallic glass thin films deposited by magnetron sputtering. *J Nanosci Nanotechnol* **5**(3): 420–424 (2005).

[14] Schroers J. Processing of bulk metallic glass. *Adv Mater* **22**(14): 1566–1597 (2010)

[15] Caron A, Sharma P, Shluger A, Fecht H-J, Louzguine-Luzguin D V, Inoue A. Effect of surface oxidation on the nm-scale wear behavior of a metallic glass. *J Appl Phys* **109**(8): 083515 (2011)

[16] Caron A, Louzguine-Luzguin D V, Bennewitz R. Structure vs chemistry: Friction and wear of Pt-based metallic surfaces. *ACS Appl Mater Interf* **5**(21): 11341–11347 (2013)

[17] Suzuki S, Osaki H, Ando E. Materialistic difference in macroscopic friction coefficients of sputtered oxide thin films deposited on glass. *Jpn J Appl Phys* **35**(3): 1862–1867 (1996)

[18] Wang D, Tan H, Li Y. Multiple maxima of GFA in three adjacent eutectics in Zr-Cu-Al alloy system–A metallographic way t pinpont the best glass forming alloys. *Acta Mater* **53**(10): 2969–2979 (2005)

[19] Inoue A, Negishi T, Kimura H M, Zhang T, Yavari A R. *Mater Trans* **39**(2): 318–321 (1998)

[20] Meyer E, Overney R M, Dransfeld K, Gyalog T. *Friction and Rheology on the Nanometer Scale*. Singapore: World Scientific, 1998.

[21] Caron A, Qin C L, Gu L, González S, ShlugerA, Fecht H-J, Louzguine-Luzguin D V, Inoue A. Structure and nano-mechanical characterisitcs of surface oxide ayers on a metallic glass. *Nanotechnology* **22**(9): 095704 (2011)

[22] Louzguine-Luzgin D V, Chen C L, Lin L Y, Wang Z C, Ketov S V, Miyama M J, Trifonov A S, Lubenchenko A V, Ikuhara Y. Bulk metallic glass surface native oxide: Its atomic structure, growth rate and electrical properties. *Acta Mater* **97**: 282−290 (2015)

[23] Bowden FP, Tabor D. *The Friction and Lubrication of Solids*. Oxford (UK): Oxford University Press, 1950.

[24] Louzguine-Luzgin D V, Nguyen H K, Nakajima K, Ketov S V, Trifonov A S. A study of the nanscale and atomic scale wear resistance of metallic glasses. *Materials Lett* **185**: 54−58 (2016)

[25] Rabinowicz E. *Friction and Wear of Materials*, 2nd Edition. New York (USA): John Wiley & Sons Inc., 1995.

[26] Zhang W, Ji G, Bu A, Zhang B. Tailoring the valence band offset of Al_2O_3 on epitaxial $GaAs_{1-y}Sb_y$ with tunable antimony composition. *ACS Appl Mat Interf* **7**(51): 28624−28631 (2015)

[27] Tocha E, Schoenherr H, Vansco G J. Influence of grain size and humidity on the nanotribological properties of wear-resistant nanostructured ZrO_2 coating: An atomic force microscopy study. *J Am Ceram Soc* **88**(9): 2498–2503 (2005)

[28] Tocha E, Pasaribu H R, Schipper D J, Schoenherr H, Vansco G J. Low friction in CuO-dopped yttria-stabilized tetragonal zirconia ceramics: A complementary macro- and nanotribology study. *J Am Ceram Soc* **91**(5): 1646–1652 (2008)

S. J. KANG. He received his bachelor degree in material science and engineering from Koreatech (Korea University of Technology and Education), Cheonan, Korea. During his bachelor program, he has done an internship program as an assistant researcher in Nano-tribology group at INM-Leibniz Institute for New Materials, Saarbrücken, Germany. He is studying for his master degree in material science and engineering at Zhejiang University. His research interests include new-structured materials, nano-tribology and mechanical properties of metallic glasses.

A. CARON. He is a materials scientist with expertise in the multi-scale mechanical behavior of materials, surfaces and micro-components. Since 2015 Arnaud Caron is assistant professor in the School of Energy, Materials and Chemical Engineering at KoreaTech–Korea University of Technology and Education, Republic of Korea. Arnaud Caron obtained his engineering degree in materials science in 2004 from the University of Saarland, Germany and was awarded with the Schiebold Medal. In 2009 he earned his doctoral degree in materials science from the University of Saarland, Germany. From 2007 to 2015 Arnaud Caron worked as a research associate at the Institute of micro- and nanomaterials of the University of Ulm, Germany, the WPI-Advanced Institute of Materials Research, Japan and the Leibniz–Institute for New Materials, Germany.