SURROGATE BASED MODELING OF JOINT CONTACT MECHANICS

By

YI-CHUNG LIN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2008

To my parents

ACKNOWLEDGMENTS

First, I would like to sincerely thank Dr. Fregly and Dr. Haftka for the direction they have given me. Their valuable suggestions and dedications contributed greatly. I would also like to thank Dr. Banks, Dr. Kim, and Dr. Jiang for serving on my doctoral committee. It was a great honor to have them. I would also like to thank all the Computational Biomechanics Lab members for their support in my past and present work. Last but not least, I thank my parents and my wife Hsiu-Jen for being so supportive.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

SURROGATE BASED MODELING OF JOINT CONTACT MECHANICS

By

Yi-Chung Lin

December 2008

Chair: Benjamin J. Fregly
Cochair: Raphael T. Haftka
Major: Mechanical Engineering

Musculoskeletal computer models are useful for estimating internal quantities that cannot be measured experimentally, designing new medical devices and rehabilitation approaches, and predicting the outcome of surgical procedures. Unfortunately, use of articular contact in such models makes computational speed a limiting factor, rendering dynamic simulations either completely intractable or else so slow that optimization is impossible. Lacking of articular contact will significantly affect many contact quantities predicted from musculoskeletal computer models (e.g., muscle forces, ligament strains, bone loads, and cartilage and implant wear).

Computational limitations in other engineering disciplines have been overcome through the use of surrogate modeling. Surrogate-based modeling involves fitting a model to a model, where the original model is computationally expensive and the surrogate model is computationally cheap. The sample data points used to fit the surrogate model are generated by running the original model repeatedly with different combinations of input variables. Once fitted, the surrogate model can be used in place of the original model in simulations or optimizations to eliminate computational cost as a limiting factor. Though surrogate modeling techniques have

successfully eliminated computational bottlenecks in other fields, they have not yet been applied to articular contact problems.

Primary objectives were two-fold. First, we present a computational evaluation and practical application of the proposed surrogate-based modeling approach using dynamic wear simulation of a total knee replacement constrained to both two- and three-dimensional motions in a Stanmore machine. The sample points needed for surrogate modeling fitting are generated by an elastic foundation contact model. Second, we apply the surrogate contact model into the musculoskeletal model to simultaneously calculate medial and lateral tibiofemoral contact forces, patellofemoral contact forces, and the muscle forces. This project provides a computationally efficient way to apply different types of contact models to predict physiologically significant contact quantities.

CHAPTER 1
INTRODUCTION

Need for Efficient Contact Model Evaluation

Musculoskeletal computer models are useful for estimating physiological quantities that cannot be measured experimentally (Buchanan and Shreeve, 1996), designing new medical devices and rehabilitation approaches (Neptune et al., 2000), and predicting the outcome of surgical procedures (Delp et al., 1996). In such models, muscles, ligaments, and bones interact (as in the real musculoskeletal system) to determine the motions and loads experienced by the anatomic structures. Much of this interaction occurs at the joints, where forces from each type of structure influence the forces exerted on and by the other structures.

Lack of articular (i.e., surface-surface) contact in musculoskeletal computer models can lead to inaccurate prediction of quantities influenced by muscle, ligament, and bone interactions. At least four types of predictions can be adversely affected by omission of articular contact models: 1) Muscle force predictions – the amount of constraint present in traditional engineering joint models (e.g., hinge versus ball and socket) can significantly influence muscle force predictions during optimization analysis (Challis and Kerwin, 1993; Glitsch and Baumann, 1997; Li et al., 1999; Pierce and Li, 2005). 2) Ligament strain predictions –articular geometry, via its influence on joint kinematics, can have a significant effect on the strains calculated in individual ligaments (Blankevoort et al., 1991; Wilson et al., 1998). 3) Bone remodeling predictions – bone remodeling simulations are sensitive to the applied joint contact and muscle forces, both of which are influenced by articular contact conditions (Bitsakos et al., 2005). 4) Cartilage and implant wear predictions – contact forces, pressures, and sliding conditions within a joint are dependent on articular surface geometry as well as muscle co-contractions and ligament forces (Kwak et al., 2000; Fregly et al., 2005).

Unfortunately, use of articular contact in musculoskeletal computer models makes computational speed a limiting factor, rendering dynamic simulations or optimization analysis too slow to perform. Unlike engineering joint models, articular contact models require repeated evaluation of surface geometry, and these geometry evaluations consume the vast majority of the CPU time in a dynamic contact simulation (Bei and Fregly, 2004). A slow dynamic simulation requiring hours or days of CPU time (Giddings et al., 2001; Godest et al., 2002) means painfully slow repeated analyses and impractical optimization studies, which typically require hundreds or even thousands of repeated analyses. While several labs have developed their own articular contact models to address computational speed issues (Blankevoort et al., 1991; Pandy et al., 1997; Kwak et al., 2000; Cohen et al., 2001; Piazza and Delp, 2001; Chao, 2003; Elias et al., 2003; Bei and Fregly, 2004; Caruntu and Hefzy, 2004), only three models have been used to perform dynamic simulations (Piazza and Delp, 2001; Caruntu and Hefzy, 2004; Bei and Fregly, 2004), only one functions within a general dynamic simulation environment (Bei and Fregly, 2004), and none can be labeled as "fast" for performing repeated dynamic simulations involving articular contact at one or more joints.

Similar computational challenges are being tackled in other engineering fields using surrogate-based modeling (Giunta et al., 1997; Kurtaran et al., 2001; Queipo et al., 2002). This approach involves fitting a computationally cheap model to data points generated by a computationally expensive model. The surrogate model (also sometimes called a "response surface") is then used in place of the original model to eliminate a computational bottleneck either in some aspect of a simulation (e.g., the contact calculations) or in an entire simulation (e.g., by providing fast evaluation of the cost function and constraints).

Unlike traditional engineering applications, joint contact analyses pose unique challenges to existing surrogate-based modeling approaches. Sampling data points within a hypercube defined by the upper and lower bounds of the relative pose parameters will result in few physically realistic data points for surrogate model creation. Most sample points will be either out of contact (i.e., zero contact force) or in excessive penetration (i.e., unrealistically high contact force) since contact forces are highly sensitive to the small displacements in the contact normal direction. Consequently, traditional surrogate-based modeling approaches fail to provide the accuracy needed to perform dynamic simulations. A special surrogate-based modeling approach adapted to the specific needs of joint contact analyses is therefore required.

Specific Aims

In short, the goal of present work is to develop and evaluate a novel surrogate contact modeling approach to permit rapid dynamic simulation, sensitivity analysis, and optimization of musculoskeletal models with articular contact. The objectives of the current work can be summarized as follows:

1. Evaluate the effect of variations in component placements and input profiles on wear volume generated by a knee simulator machine.
2. Develop and evaluate an extended surrogate contact modeling approach capable of simulating three-dimensional situations.
3. Predict muscle and contact forces simultaneously at the knee using a musculoskeletal model incorporating three-dimensional surrogate contact models.
4. Evaluate the muscle and contact forces prediction using in vivo contact force measurements provided by an instrumented knee replacement.

CHAPTER 2
TWO-DIMENSIONAL SURROGATE CONTACT MODELING FOR COMPUTATIONALLY-EFFICIENT DYNAMIC SIMULATION OF TOTAL KNEE REPLACEMENTS

Introduction

For more than three decades, total knee replacement (TKR) surgery has been performed on patients with severe osteoarthritis. Though modern TKR surgery has a high success rate, some patients still experience substantial functional impairment post-surgery compared to healthy individuals (Noble et al., 2005). Furthermore, patients today desire more than just pain relief, seeking a high level and broad variety of daily activities following TKR surgery (e.g., tennis, hiking, gardening, even jogging) (Mancuso et al., 2001; Weiss et al., 2002). Thus, maximization of durability and minimization of functional limitations have become two primary design goals.

One way to pursue these goals is through the development of computational technology that permits sensitivity and optimization studies of new TKR designs. A recent step in this direction has been the development of computational models of knee wear simulator machines (Giddings et al., 2001; Halloran et al., 2005a; Rawlinson et al., 2006; Fregly et al., 2008; Zhao et al., 2008). Such models permit dynamic contact and wear simulations of knee implant designs to be performed in a time- and cost-efficient manner, with model predictions being validated against experimental motion and wear measurements (Halloran et al., 2005a; Knight et al., 2007; Fregly et al., 2008; Zhao et al., 2008). Though finite element models can be used to perform contact calculations for these simulations, recent work has shown that elastic foundation contact models produce similar contact forces and pressures in a fraction of the CPU time (Halloran et al., 2005b). Even so, CPU times on the order of 5 to 10 minutes for a one-cycle dynamic contact simulation can still be prohibitive for design sensitivity and optimization studies requiring thousands of simulations.

Computational limitations in other engineering disciplines have been overcome through the use of surrogate modeling approaches. Surrogate modeling (also known as "metamodeling" or "response surface approximation") involves replacing a computationally costly model with a computationally cheap model constructed using data points sampled from the original model. Once the surrogate model is constructed, it is used in place of the original computationally costly model when subsequent engineering analyses are performed. A variety of surrogate model fitting methods have been proposed, including polynomial response surfaces (Box and Draper, 1987; Myers and Montgomery, 1995; Khuri and Cornell, 1996), Kriging (Sacks et al., 1989; Jones et al., 1998), radial basis functions (Wendland, 1995; Chen et al., 1996), splines (Wahba, 1987), and support vector machines (Girosi, 1998; Vapnik, 1998). While surrogate models have been utilized successfully for a variety of engineering applications (Koch et al., 1999; Liu et al., 2000; Cox et al., 2001; Queipo et al., 2002; Queipo et al., 2005; Wang and Shan, 2007), only a few studies have used surrogate models to fit input-output relationships from a computational contact model (Bouzid et al., 1998; Chang et al., 1999; Lin et al., 2006). Furthermore, none of these studies has proposed a surrogate modeling technique to improve the computational efficiency of dynamic contact simulations.

This study proposes a surrogate contact modeling approach for performing dynamic contact and wear simulations of total knee replacements in a computationally efficient manner. The approach addresses the unique challenges involved in applying surrogate modeling techniques to joint contact problems. The primary objectives of this chapter are two fold. The first is to evaluate the computational speed and accuracy of the proposed surrogate contact modeling approach by comparing its results with those generated by an elastic foundation contact model. Sample points for the surrogate contact model are generated by the same elastic

foundation contact model used in the comparative dynamic wear simulations. The second objective is to demonstrate the practical applicability of the proposed approach by analyzing the sensitivity of knee replacement wear predictions to realistic variations in input motions and loads and component placements. A Monte Carlo approach requiring thousands of dynamic wear simulations is used to perform the sensitivity study. Both objectives are pursued using a dynamic contact model of a total knee replacement constrained to planar motion in a Stanmore machine, and both demonstrate the significant computational benefits that surrogate contact modeling can provide for analysis of TKR designs.

Methods

Surrogate Contact Model Development

We have developed a novel surrogate modeling approach to replace computationally costly contact calculations in dynamic simulations of total knee replacements. To develop the approach, we utilized a three-dimensional (3D) elastic foundation contact model (An et al., 1990; Blankevoort et al., 1991; Li et al., 1997; Pandy et al., 1998; Fregly et al., 2003) of a cruciate-retaining commercial knee implant (Osteonics 7000, Stryker Howmedica Osteonics, Inc, Allendale, NJ) constrained to sagittal plane motion in a Stanmore simulator machine (Walker et al., 1997). The model possessed three degrees of freedom (DOFs) relative to ground: tibial anterior-posterior translation, femoral superior-inferior translation, and femoral flexion-extension. Similar to a Stanmore machine, femoral flexion-extension rotation was prescribed while the remaining two DOFs were loaded using ISO standard input curves (DesJardins et al., 2000). A pair of parallel spring bumper loads was also applied to the tibia in the anterior-posterior direction to approximate the effect of ligament forces (Walker et al., 1997). The EF contact model was treated as an applied load on the femoral component and tibial insert and used linear elastic material properties (Fregly et al., 2005). The dynamic equations for the

model were derived via Kane's method using Autolev symbolic manipulation software (OnLine Dynamics, Sunnyvale, CA). The equations were incorporated into a Matlab program (The Mathworks, Natick, MA) that was used to perform forward dynamic simulations with the stiff solver ode15s.

To develop a surrogate contact model to replace the EF contact model in the larger dynamic model described above, we followed a four-step process: 1) design of experiments (DOE), 2) computational experiments, 3) surrogate model selection, and 4) surrogate model implementation (Figure 2-1). The first step, DOE, is a statistical method for determining at which locations in design space computational experiments should be performed with the computationally costly model to sample the outputs of interest (Figure 2-1A). For our sagittal plane knee implant model, locations in design space (or "sample points") were defined by the following three relative pose parameters: tibial anterior-posterior (AP) translation X, femoral superior-inferior (SI) translation Y, and femoral flexion-extension rotation θ. The associated outputs of interest calculated by the EF contact model were the AP contact force F_x and the SI contact force F_y. To minimize the total number of sample points while maximizing the quality of the resulting surrogate model fit, we chose a Hammersley quasirandom (HQ) sampling method (Hammersley, 1960). This method uses an optimal design scheme to uniformly place sample points within a multi-dimensional hypercube (Kalagnanam and Diwekar, 1997; Diwekar, 2003).

The second step, computational experiments, involved performing repeated simulations with the original contact model at the sample points defined in the first step. Because contact forces are highly sensitive to some pose parameters but not others, we made one modification to the computational experiments performed at the HQ sample points (Figure 2-1B). A traditional HQ sampling method would sample values of the three inputs X, Y, and θ within their specified

bounds and use the EF contact model to calculate the two outputs F_x and F_y at each sample point. However, since SI contact force F_y is highly sensitive to small variations in SI translation Y (Lin et al., 2006; Fregly et al., 2008), this approach produced a large number of sample points that were either out of contact or deeply penetrating. Inclusion of these unrealistic points in the surrogate model fitting process reduced the accuracy of the resulting surrogate contact model. To resolve this issue, we inverted the HQ sampling method for superior-inferior translation by sampling forces in the sensitive direction (i.e., F_y) and displacements in the insensitive directions (i.e., X and θ). Specifically, we used the HQ method to sample 300 pairs of X and θ values. For each of these pairs, we generated five sample points by performing five static analyses with the EF contact model using F_y values ranging from 600 to 2400 N (approximately 3x body weight) in 600 N increments plus a 10 N value for determining contact initiation. The outputs of each static analysis were F_x and Y. Evaluating 1500 sample points (X, F_y, θ) using static analyses required approximately 3 hours of CPU time on a 3.4 GHz Pentium IV PC.

The third step, surrogate model selection, involved determining the most appropriate type of surrogate model for fitting the input-output relationships defined by the computational experiments in the second step. Based on its ability to interpolate multi-dimensional non-uniformly spaced sample points, Kriging (Sacks et al., 1989; Cressie, 1992) was selected as the surrogate modeling method to fit the input-output relationships defined by the 1500 sample points. Originally developed for geostatistics and spatial statistics, Kriging has been widely applied in many different fields (Chang et al., 1999; Gupta et al., 2006). In mathematical terms, Kriging utilizes a combination of a polynomial trend model and a systematic departure model:

$$y(\mathbf{x}) = f(\mathbf{x}) + z(\mathbf{x}) \quad (2\text{-}1)$$

where $\mathbf{x}$ is a vector of design variable inputs, $y(\mathbf{x})$ is the output to be fitted, $f(\mathbf{x})$ is a polynomial trend model, and $z(\mathbf{x})$ is a systematic departure model whose correlation structure is a function of

the distances between sample points. The f(**x**) term approximates the design space globally while the z(**x**) term interpolates local deviations from f(**x**). The necessary parameters for z(**x**) were obtained through maximum likelihood estimation (Sacks et al., 1989). The Matlab toolbox DACE (Lophaven et al., 2002) was used to construct all necessary Kriging models using a cubic polynomial trend function and a cubic spline correlation function based on preliminary tests.

The fourth step, surrogate model implementation, required the development of a special model inversion process to accommodate our modified HQ sampling approach (Figure 2-1C). During a dynamic simulation, the EF contact model requires X, Y, and θ as inputs and provides F_x and F_y as outputs, whereas the surrogate contact model requires X, F_y, and θ as inputs and predicts F_x and Y as outputs. To address this inconsistency, we fitted F_y as a function of Y using the general form of a Hertzian contact model (Equation 2-2).

$$\hat{F}_y = k_y(X,\theta)\left[Y_{10N}(X,\theta) - Y\right]^{n_y(X,\theta)} \tag{2-2}$$

In this equation, $\hat{F}_y$ is the predicted SI contact force, k_y is the contact stiffness, Y_{10N} is the SI translation at contact initiation, and n_y is the contact exponent. Use of Equation 2-2 to calculate $\hat{F}_y$ required fitting k_y, Y_{10N}, and n_y as a function of X and θ. Each of the 300 (X, θ) sample pairs had five sampled values of F_y (i.e., 10N, 600N, 1200N, 1800N, 2400N) and five associated output values of Y. Thus, Y_{10N} was already known for each (X, θ) sample pair. To find k_y and n_y for each (X, θ) sample pair, we linearized Equation 2-2 by taking $\log_{10}$ of both sides and solved for $\log_{10}(k_y)$ (and hence k_y) and n_y using linear least squares (5 equations in 2 unknowns). This process yielded one value of k_y, Y_{10N}, and n_y for each (X, θ) sample pair.

To calculate the AP contact force $\hat{F}_x$, we used Equation 2-3 based on the observation that F_x is close to a constant fraction of F_y at each (X, θ) sample pair. This fraction (*ratio*) is the

average of the five (F_x, F_y) quotients obtained for each sample pair, and $\hat{F}_y$ is the SI contact force predicted by Equation 2-2. Thus, the final surrogate contact model utilized four separate Kriging models to fit k_y, Y_{10N}, n_y, and *ratio* as a function of X and θ (Figure 2-1D), permitting the calculation of $\hat{F}_y = f(X, Y, \theta)$ and $\hat{F}_x = f(X, \theta, \hat{F}_y)$ from Equation 2-2 and 2-3 at any point during a dynamic simulation (Figure 2-1E).

$$\hat{F}_x = ratio\left(X, \theta\right) \hat{F}_y \tag{2-3}$$

Surrogate Contact Model Evaluation

To evaluate the computational speed and accuracy of our surrogate contact modeling approach, we performed nine dynamic wear simulations with both the surrogate contact model and the EF contact model used to generate it. Each wear simulation utilized different nominal component placements that were variations away from the original placements (Figure 2-2A). The specific variations simulated were all possible combinations of three SI locations of the femoral flexion axis (original ± 10 mm) and three AP slopes of the tibial insert (original ± 10 deg). To cover all possible dynamic simulation paths, we specified the ranges of the 300 HQS points using the limits of motion obtained from dynamic simulations performed using the EF contact model when the SI locations of the femoral flexion axis and AP slope of the tibial insert were set to their extreme values. We quantified differences between the sets of results by calculating root mean square errors (RMSE) in predicted contact forces, component translations, and wear volumes.

Generation of wear volume predictions with the surrogate contact model required the development of a separate surrogate modeling approach based on prediction of medial and lateral center of pressure (CoP) locations. This approach used Archard's wear law (Archard and Hirst, 1956, and Appendix A) to predict wear volume for any specified number of cycles from the time

history of contact forces and CoP slip velocities on each side over a one-cycle dynamic simulation (Equation 2-4).

$$V = N\sum_{i=1}^{n}\sum_{j=1}^{2} kF_{ij}d_{ij} = Nk\sum_{i=1}^{n}\sum_{j=1}^{2} F_{ij}\left|\mathbf{v}_{ij}\right|\Delta t \tag{2-4}$$

In this equation, V is the total wear volume of the medial and lateral sides over N cycles, N is the assumed total number of cycles (5 million), i is a discrete time instant during the one-cycle dynamic simulation (1 through n), j is the side (medial or lateral), k is the assumed material wear rate (1×10^{-7} mm^3/Nm; Fisher et al., 1994), F_{ij} is the total contact force at time instant i on side j, and d_{ij} is the sliding distance of the CoP at time instant i on side j. Due to symmetry of the implant geometry, the predicted value of F_{ij} on each side was taken as half the total contact force obtained from Equation 2-2 and 2-3. Prediction of d_{ij} on each side required determining the magnitude of the slip velocity $\mathbf{v}_{ij}$ of the CoP and multiplying by the time increment Δt (Fregly et al., 2005). To find $\mathbf{v}_{ij}$, we used the kinematic concept of two points fixed on a rigid body (Equation 2-5).

$$\mathbf{v}_{ij} = {}^{A}\mathbf{v}^{O} + {}^{A}\boldsymbol{\omega}^{B} \times \mathbf{p}^{OCoP} \tag{2-5}$$

In this equation, A represents the tibial insert, B represents the femoral component, O is the femoral coordinate system origin located at the intersection of the flexion-extension and superior-inferior translation axes, CoP is the medial or lateral center of pressure location treated as a point fixed in body B, ${}^{A}\mathbf{v}^{O}$ is the velocity of point O with respect to reference A as determined from the superior-inferior translation of the femoral component, ${}^{A}\boldsymbol{\omega}^{B}$ is the angular velocity of body B with respect to reference frame A as determined from the flexion-extension of the femoral component, and $\mathbf{p}^{OCoP}$ is the position vector from point O to point CoP. While the values of ${}^{A}\mathbf{v}^{O}$ and ${}^{A}\boldsymbol{\omega}^{B}$ were obtained directly from the dynamic simulations, calculation of

the CoP location on each side required six additional Kriging models (i.e., 3 components per side). Each Kriging model was created by fitting the average of the 5 CoP locations obtained from each (X, θ) sample pair. Wear volumes calculated by the surrogate contact model using the CoP-based approach were compared with those calculated by the EF contact model using both the CoP-based approach and a previously published element-based approach (Fregly et al., 2005; Zhao et al., 2008) to verify that the surrogate model wear volume predictions were accurate.

Surrogate Contact Model Application

As a practical application of the proposed surrogate contact modeling approach, we performed two types of sensitivity analyses to investigate how machine set-up issues affect predicted wear volume. The first was a global sensitivity analysis to investigate the effect of large variations in component placements within the ranges defined in Figure 2-2A (i.e., ±10 mm/deg). Specifically, we generated 400 combinations of component placements using 20 evenly spaced values of femoral flexion axis SI location and tibial insert AP slope. For each combination, the surrogate-based contact model was used to perform a dynamic simulation and predict wear volume. Percent changes in predicted wear volume were calculated relative to the maximum predicted value from all global sensitivity analysis results. The second category was a local sensitivity analysis to investigate the effect of small variations in nominal component placements and motion and load inputs by using Monte Carlo techniques (Figure 2-2B). For the local sensitivity analysis, allowable variations in motion and load input curves were defined using experimental observations from eight different implant designs tested in a Stanmore simulator machine (DesJardins et al., 2000). Deviations of the experimental input curves from the ISO standard curves were calculated and analyzed using principal component analysis (Daffertshofer et al., 2004), with the first two principal components being sufficient to account

for over 98% of the variability in each type of curve. These two components were used to generate new deviation curves by selecting component weights as uniformly distributed random numbers within the bounds of the experimental curves. These deviation curves were added to the ISO standard curves to create new input motion and load curves as needed for the Monte Carlo analyses.

To distinguish between the effects of small errors in motion and load inputs and small errors in component placements, the local sensitivity analysis performed four types of Monte Carlo analyses for each of the nine nominal component placements for a total of 36 Monte Carlo analyses in all. Each type of Monte Carlo analysis varied input profiles and component placements separately or together. The first type varied motion and load profiles within 100% of their allowable ranges and also varied component placements within ± 1 mm/deg. The second type was similar but used smaller variations of only 10% of the allowable range for motion and load profiles and ± 0.1 mm/deg for component placements. The third type varied only motion and load profiles within 100% of their allowable ranges and imposed no variations on component placements. Finally, the fourth type varied component placements within ± 1 mm/deg but imposed no variations on motion and load profiles. For all types, percent changes in predicted wear volume were calculated relative to values obtained from the nominal component placements using the ISO standard inputs. Each Monte Carlo analysis was performed on a 3.4 GHz Pentium IV PC and involved at least 1000 dynamic contact simulations. The convergence criterion for each analysis was met when the mean and coefficient of variance (i.e., 100*standard deviation/mean) for the last 10% of the wear predictions were within 2% of the final mean and coefficient of variance (Valero-Cuevas et al., 2003).

Results

For the computational evaluation, the dynamic simulations performed with the surrogate contact model closely reproduced the planar motions, contact forces, and wear volumes predicted with the EF contact model for all nine component placements. On average, the RMSE for planar motions and contact forces were less than 0.125 mm and 4 N, respectively (Figure 2-3). While each dynamic simulation performed with the EF contact model required approximately 13 minutes of CPU time, each simulation performed using the surrogate contact model required 13 seconds. For both the surrogate contact model and the EF contact model, the CoP-based approach for predicting wear volumes reproduced the element-based EF results to within 0.5% error (Table 2-1).

For the practical application, the global and local sensitivity analyses revealed that large variations in component placements relative to the original locations and small variations in input motion relative to the ISO standard had a significant affect on predicted wear volume. For the global sensitivity analysis, predicted wear volume was sensitive to large variations in the SI location of the femoral flexion axis (maximum change of 26% when AP slope of tibial insert held fixed), the AP slope of the tibial insert (maximum change of 13% when SI location of femoral flexion axis held fixed), or both parameters simultaneously (maximum change of 33%) (Figure 2-4). For the local sensitivity analysis, wear volume ranges were larger (mean of 18 mm^3 compared to 2 mm^3) and 50% percentile wear volume results were lower (mean of 23 mm^3 compared to 31 mm^3) when motion and load inputs and component placements were varied within 100% rather than 10% of their allowable ranges (Figure 2-5A and 2-5B). When inputs and placements were varied separately within 100% of their allowable ranges, predicted wear volume was much more sensitive to motion and load inputs than to component placements (maximum change of 64% compared to 5%; Figure 2-5C and 2-5D). Finally, when an additional global

sensitivity analysis was performed with the original component placements to separate the influence of the different inputs, predicted wear volume varied approximately linearly with small changes in each input curve but was more sensitive to changes in flexion-extension motion (maximum change of 47%; Figure 2-6) than to changes in input loads (maximum change of 20% for F_x and 14% for F_y; Figure 2-6A and B). Each Monte Carlo analysis performed with the surrogate contact model required 4 hours of CPU time compared to an estimated 230 hours with the EF contact model.

Discussion

Computational speed is a major limiting factor for performing sensitivity and optimization studies of total knee replacement designs. Unlike constraint-based joint models, knee joint contact models require repeated geometry evaluations that consume the vast majority of the CPU time in a dynamic contact simulation (Bei and Fregly, 2004). This paper has presented a surrogate modeling approach for performing efficient dynamic contact simulations of human joints. The surrogate contact model was developed to replace the original EF contact model to avoid time-consuming repeated evaluation of the surface geometry during dynamic simulation. The surrogate contact model produced dynamic simulation and wear results that accurately matched those from the EF contact model but with an order of magnitude improvement in computational speed. This speed improvement would have been even more significant if a finite element contact model had been replaced. The 36 Monte Carlo analyses demonstrated how analyses that would have been impractical (or at a minimum would have required extensive parallel processing) can be easily achieved on a single computer with the surrogate contact model. The sensitivity results also have practical value, suggesting that wear volume generated by simulator machines may be sensitive to large variations in component placements within the machine as well as small variations in the flexion-extension motion input to the machine.

The primary benefit of using surrogate contact models in dynamic simulations is improved computational efficiency. For elastic foundation or finite element contact models, the computation time per dynamic simulation is largely determined by the number of deformable contacts in the model. Thus, a musculoskeletal computer model utilizing deformable contact models for multiple joints could easily require hours or days of CPU time to complete a single dynamic simulation. Repeated dynamic simulations as part of a sensitivity or optimization study would become impractical. Although it took 3 hours of CPU time on generating 1500 sample points, the use of the surrogate contact model in place of the EF model reduced Monte Carlo analysis time from 230 hours to 4 hours, a nearly 99% reduction. By using surrogate contact models instead, one could eliminate the computational cost of the contact solver as the limiting factor. The larger the number of deformable contacts in a multibody dynamic model, the greater the anticipated benefit from using surrogate contact models.

While we have shown that surrogate contact models can be beneficial for sensitivity studies, the greatest computational benefit is likely to occur for optimization studies. For stochastic sensitivity studies, statistically-based alternatives to Monte Carlo analyses exist that require fewer repeated simulations. For example, Laz et al. (2006) used the mean value method (Wu et al., 1990) to perform probabilistic elastic foundation simulations of a TKR design and reduced the number of repeated simulations by a factor of four compared to a traditional Monte Carlo approach. Since the CPU time per simulation was 6 minutes, use of a surrogate contact model could still reduce the overall computation time by an additional factor of 28, assuming we could achieve comparable speed improvements for the 3D situation. In contrast, no good methods exist for reducing the number of repeated simulations in optimization studies. For those

situations, the factor of 60 reduction in computation time (i.e., 13 minutes to 13 seconds) could mean the difference between an impossible and an achievable optimization study.

Other than computational efficiency, surrogate contact models possess the additional benefits of being adaptable and flexible. First, the surrogate-based contact modeling approach presented here is not limited to the knee joint. Theoretically, the surrogate-based approach could be applied to any joint contact model whose input-output relationships can be sampled. The modified HQ sampling method should be suitable for any joint contact model once the sensitive directions are identified. Therefore, neither type of joint (e.g., hip or knee or ankle) nor type of contact model (e.g., EF or finite element) nor type of material model (e.g., linear or nonlinear) is critical for developing a surrogate contact model. The benefit of using an EF contact model to generate the required sample points was that it is relatively fast computationally for performing repeated static analyses. Finite element models would require more time for each static analysis, plus the inputs are usually applied displacements rather than loads. For the finite element situation, it could be necessary to curve fit the force-displacement curve at each sample point to produce new points at fixed force increments. Second, the surrogate contact model represents the real surface geometry implicitly within the design space. Since the surrogate modeling approach uses the real surface geometry to generate the sampling points, no idealization of complex surface geometry is required. Third, the approach is not limited to a single surrogate model fitting method. While we used Kriging in the present study, any of the fitting methods mentioned previously could be investigated as well (Jin et al., 2001; Simpson et al., 2001; Queipo et al., 2005; Wang and Shan, 2007). Though polynominal response surfaces are the most commonly used surrogate model, Kriging has the advantage of interpolating the sample point outputs, which is appealing for the deterministic computer models (Sacks et al., 1989; Simpson et al., 2001).

The global sensitivity results demonstrate the importance of developing a standardized method for defining the location of the femoral flexion axis. For a given tibial slope, predicted wear volume varied approximately linearly with changes in the superior-inferior location of the femoral flexion axis, increasing as the axis was translated superiorly (Figure 2-4). Superior translation of the flexion axis increases the distance to the CoP, thereby increasing the magnitude of the slip velocity in Equation 2-4. Thus, to minimize wear for any given implant design, one can simply locate the femoral flexion axis close to the articular surfaces. In contrast, the predicted wear volume varied approximately quadratically with changes in tibial slope, increasing as the insert was tilted anteriorly or posteriorly away from its neutral position (Figure 2-4). Increasing the tibial anterior or posterior slope increases AP tibial sliding distance (Bai et al., 2000), thereby leading to an increase in predicted wear volume (Blunn et al., 1991; Kawanabe et al., 2001).

In contrast, the local sensitivity results demonstrate the importance of controlling machine inputs closely, especially the femoral flexion angle. The Monte Carlos analyses and the additional global sensitivity analysis revealed when small variations were imposed on the input motion, loads, and component placements, only variations in the femoral flexion angle resulted in large changes in predicted wear volume. When the flexion angle was varied within 100% of its allowable range, the drop in 50^{th} percentile wear volume results (Figure 2-5A and 2-5C) relative to the ISO standard inputs (i.e., 50% percentile results in Figure 2-5B and 2-5D) was due to the use of imposed variations that always reduced the amplitude of the curve (DesJardins et al., 2000). This finding is consistent with previous studies showing that predicted wear volume decreases when the amplitude of the input curves is reduced (Barnett et al., 2001; D'Lima et al., 2001; Johnson et al., 2001). Thus, another way to minimize wear for any given implant design is

to use a control system that systematically undershoots the peaks in the input flexion-extension curve. Combining this approach with an inferiorly located femoral flexion axis could cut wear volume by more than 50% compared to the original component placement with ISO standard inputs. The development of a standard for acceptable variations in input flexion-extension motion should be considered as well.

Despite its computational advantages, the proposed surrogate contact modeling approach possesses several important limitations. The biggest one is that new surrogate contact models must be generated if implant geometry, insert thickness, or material properties are changed, at least for the current formulation. The geometry limitation makes it impossible to use surrogate contact models for progressive wear simulations, where the surface geometry is changed gradually over an iterative sequence of wear simulations. However, this limitation is not serious if only wear volume is of interest, as recent progressive wear simulation studies have reported that predicted wear volume (but not wear area or depth) is relatively insensitive to whether or not the surface geometry is changed gradually (Knight et al., 2007; Zhao et al., 2008). Theoretically, changes in insert thickness or material model parameter values could be accommodated by adding more design variables to the surrogate model fitting process (e.g., make insert thickness an additional surrogate model input). For contact forces, a changing value of Young's modulus (E) in a linear elastic material model can be accommodated without adding another design variable. If all sample points are generated using a value of $E = 1$, then the contact forces output by the surrogate model can be scaled by the desired value of E (Lin et al., 2006).

A number of other limitations exist as well. First, the current surrogate contact model is restricted to sagittal plane motion, where contact force is dominated by SI translation of the femoral component. For 3D simulations, contact force will be highly sensitive to changes in

varus-valgus rotation as well (Fregly et al., 2008). In addition to an increased number of sensitive directions, the number of sample points will also need to increase. However, the increased computational cost for evaluating more sample points is paid up front since the sampling process is only performed once. Second, the calculation of F_x is dependent on the value of F_y computed from Equation 2-2. Thus, the quality of the F_y prediction will directly impact the F_x calculation. The current approach works fine for sagittal plane models with a single sensitive direction, but an independent calculation of F_x may be needed for 3D situations where more than one sensitive direction is present. Third, the CoP approach for calculating wear volume may not work well for some types of 3D motion. For example, for a sphere spinning about a vertical axis while in contact with a plane, the CoP would be located at the bottom of the sphere where the slip velocity is zero. Thus, no wear volume would be calculated, even though the element-based approach would calculate a non-zero wear volume based on the pressures and slip velocities throughout the contact area. Consequently, the CoP approach will work well for analysis of TKR designs undergoing 3D motion only if the amount of internal-external rotation is "small." Fourth, our Monte Carlo results are limited to cruciate-retaining designs with a non-conformal tibial insert. Whether or not our conclusions are generalizable to other implant designs and simulator machines will require further investigation.

In summary, this chapter has demonstrated that surrogate contact modeling can significantly improve the computational speed of dynamic contact simulations and is appropriate for 2D sensitivity and optimization studies incorporating deformable contact. Refinement of the current 2D approach and extension to 3D problems are topics of ongoing research. Once surrogate contact models are developed for 3D situations, it should be possible to create dynamic musculoskeletal models with multiple deformable joints that can be simulated in a short amount

of CPU time. Such models should improve our ability to analyze new knee implant designs rapidly and to predict muscle forces across joints where contact forces may have an important stabilizing effect.

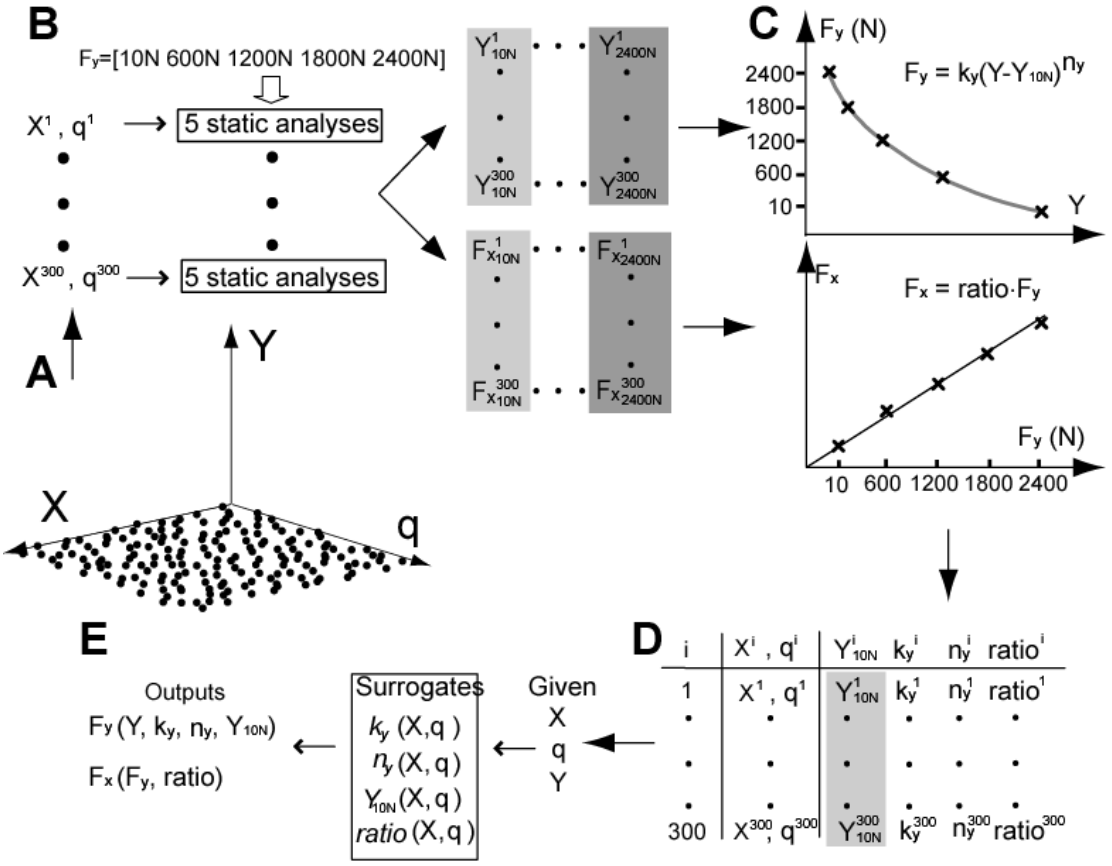

Figure 2-1. Overview of surrogate contact model creation and implementation process. A) 300 points (black dots) are sampled in the (X, θ) design space. B) Five static analyses using five axial loads are performed at each sample point location using an elastic foundation contact model. C) Generalized Hertzian contact theory is used to model F_y as a function of Y while F_x is modeled as a linear function of F_y. D) The values of Y_{10N} , k_y, n_y, and *ratio* from the 300 sample points are fitted as functions of X and θ using Kriging. E) During a dynamic simulation, the surrogate contact model calculates F_y and F_x from four Kriging models given the current values X, Y, and θ.

Table 2-1. Comparison between the wear volume predictions (mm^3) made with the elastic foundation contact model and the surrogate contact model for each of the nine nominal component placements. Each prediction was made for 5 million motion cycles. Both contact models can predict wear volume using a simple center of pressure-based approach, while only the elastic foundation contact model can make predictions using a more detailed element-based approach.

	Elastic foundation contact model		Surrogate contact model
Nominal Placement	Element-based	CoP-based	CoP based
A	37.07	37.05	36.98
B	34.03	34.02	33.91
C	39.04	39.03	38.87
D	32.91	32.89	32.88
E	30.78	30.76	30.77
F	34.20	34.19	34.14
G	29.37	29.36	29.32
H	26.45	26.44	26.38
I	29.37	29.35	29.31

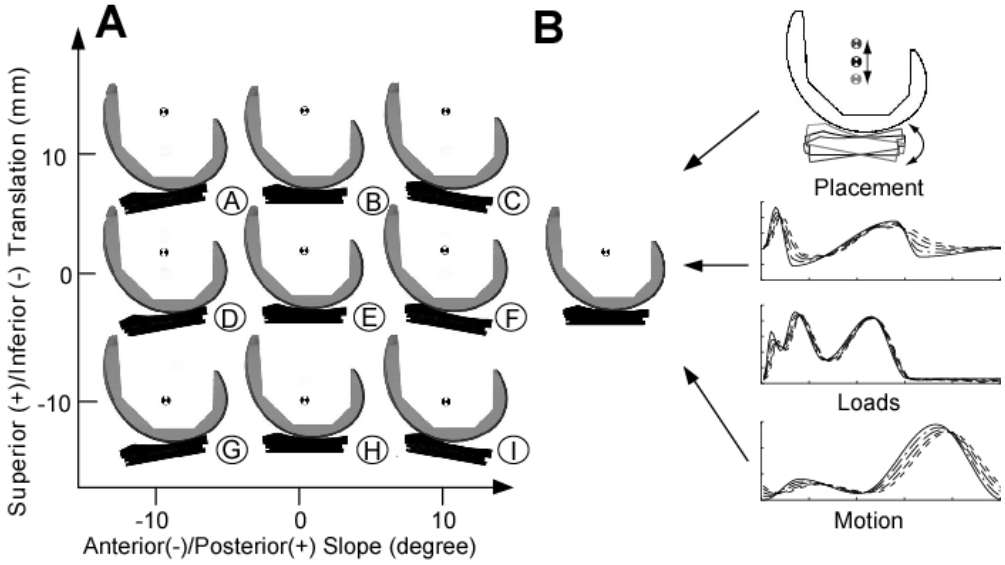

Figure 2-2. Overview of global and local sensitivity analyses of predicted wear volume performed with the surrogate contact model. A) Large variations in nominal component placements (labeled A through I) used for the global sensitivity analysis. B) Small variations in component placements, input loads, and input motion applied about each nominal component placement used for the local sensitivity analysis involving Monte Carlo methods

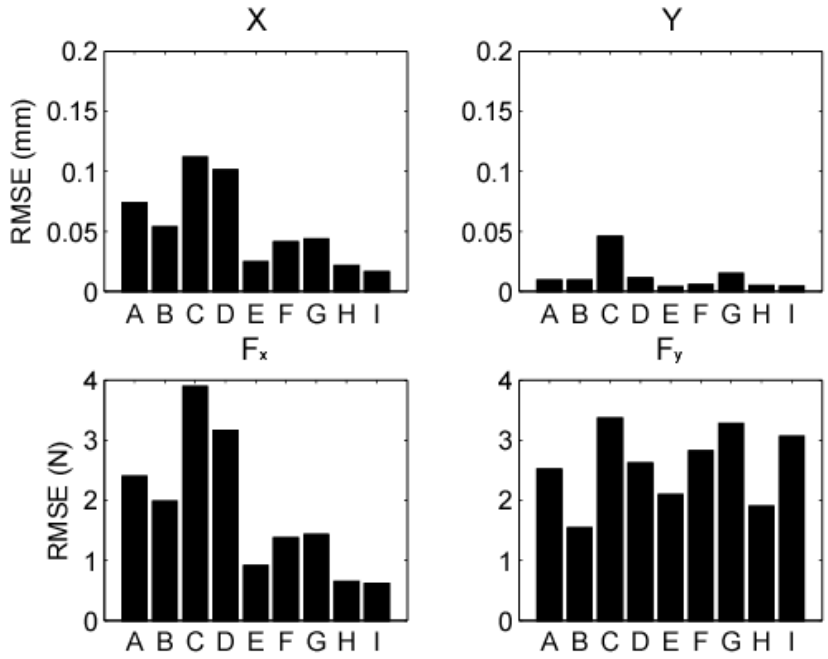

Figure 2-3. Root-mean-square errors in joint motions and contact forces for each nominal component placement. Errors are computed by comparing dynamic simulation results generated with the elastic foundation contact model and the surrogate contact model.

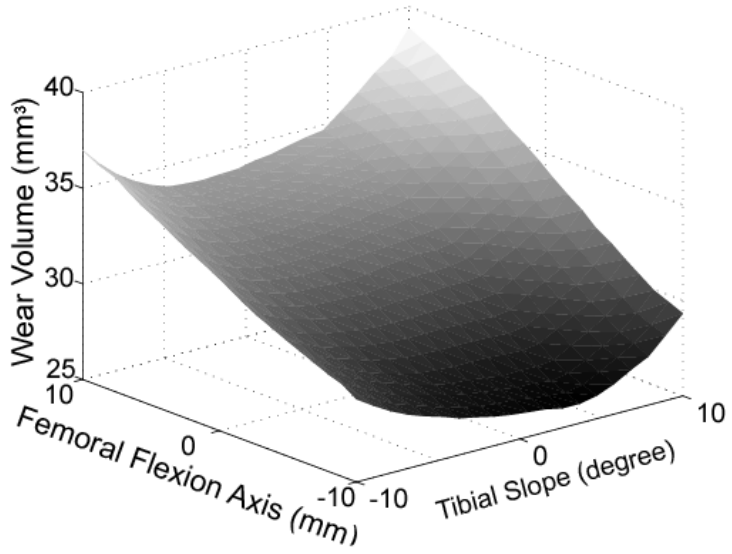

Figure 2-4. Predicted wear volume as a function of large variations in component placements as calculated from 400 dynamic simulations performed with the surrogate contact model. Wear increases linearly when the femoral flexion axis is translated superiorly and quadratically when the tibial slope is increased anteriorly or posteriorly.

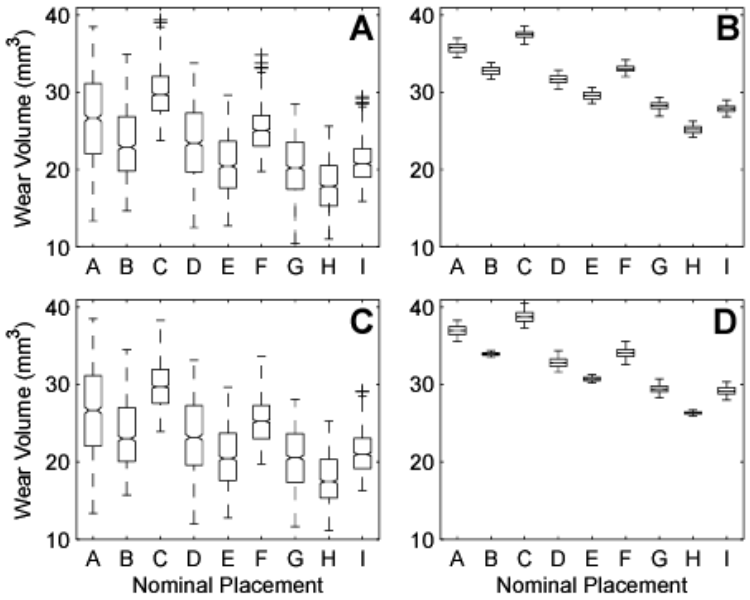

Figure 2-5. Box plot distribution of predicted wear volume generated by four Monte Carlo analyses using the surrogate contact model. Each box has (from bottom to top) one whisker at the 10th percentile, three lines at the 25th, 50th, and 75th percentile, and another whisker at the 90th percentile. Outliers are indicated by black crosses located beyond the ends of the whiskers. For the first and second Monte Carlo analyses, motion and load inputs and component placements are varied together within A) 100% or B) 10% of their maximum ranges, respectively. For the third and fourth Monte Carlo analyses, C) motion and load inputs or D) component placements are varied separately within 100% of their maximum ranges, respectively.

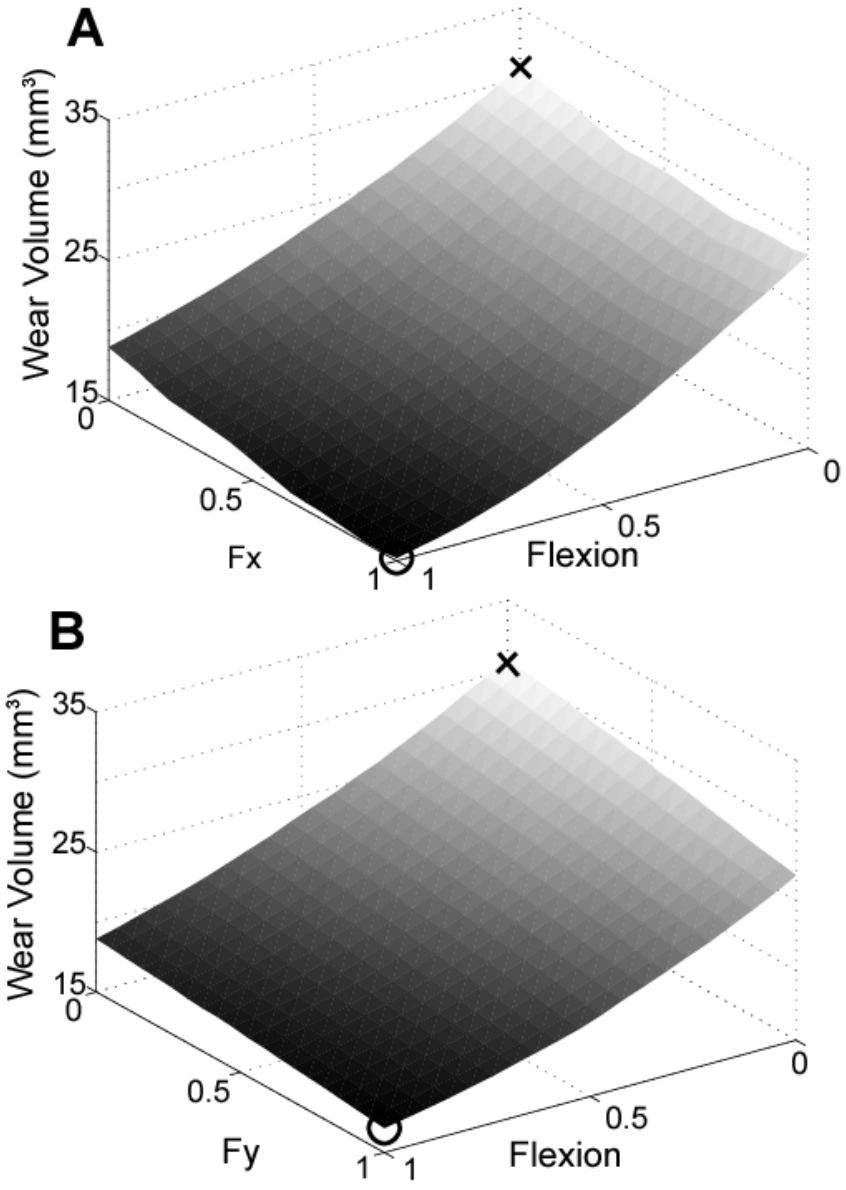

Figure 2-6. Predicted wear volume as a function of variations in A) anterior-posterior load and femoral flexion angle and B) axial load and femoral flexion angle. Both plots were created by performing dynamic simulations with the surrogate model and changing the weight of the first principal component for each input profile variation away from the ISO standard. Each weight was normalized to be between 0 and 1. The cross marker (x) represents the ISO standard input with no superimposed input profile variations (i.e., weights of 0) while the circle marker (o) represents the ISO standard input with maximum superimposed input profile variations (i.e., weights of 1)

CHAPTER 3
THREE-DIMENSIONAL SURROGATE CONTACT MODELS FOR COMPUTATIONALLY-EFFICIENT MULTIBODY DYNAMIC SIMULATIONS

Introduction

Multibody dynamic simulation of systems experiencing one or more contacts is valuable for engineering applications ranging from the design of industrial machines and mechanisms to the analysis of human joints. Computational contact models used to perform such simulations generally fall into two categories: rigid body contact models and deformable body contact models (Sharf and Zhang, 2006). Rigid body contact models use unilateral constraints to maintain contact between opposing surfaces and are highly efficient for predicting motion (Glocker, 1999; Piazza and Delp, 1999; Stewart, 2000; Porta et al., 2005; Brogliato et al., 2002). However, they do not permit calculation of unique contact forces under statically indeterminate conditions, such as when contact occurs in two or more places between the same pair of contacting bodies (Cheng et al. 1990). Furthermore, they normally require different computational schemes to simulate continuous contact versus intermittent impact.

In contrast, deformable body contact models, such as finite element models (Géradin and Cardona, 2001; Halloran et al., 2005a; Gerstmayr and Schoberl, 2006) and elastic foundation models (Fregly et al., 2003; Hippmann et al., 2004; Santamaria et al., 2006; Moran et al., 2008), provide a unified approach for simulating continuous contact versus intermittent impact and can calculate unique contact forces in statically indeterminate situations. Furthermore, they permit calculation of distributed contact pressures and strains across the opposing surfaces. These benefits are achieved by using elastic contact theory rather than constraints to determine the forces and torques exerted by the contacting bodies on each other. Unfortunately, the high computational cost of deformable body contact analyses, particularly due to extensive geometry evaluations, is a significant limiting factor to their use in multibody dynamic simulations (Bei

and Fregly, 2004). Thus, no contact modeling approach currently exists that provides the benefits of deformable body contact methods and the computational speed of rigid body contact methods.

In other engineering disciplines, surrogate modeling approaches have been used successfully to overcome similar computational challenges (e.g., Roux et al., 1996; Liu et al., 2000; Cox et al., 2001; Queipo et al., 2002; Wang et al., 2007). In general terms, these approaches involve replacing a computationally costly model with a computationally cheap model constructed using data points sampled from the original model. Once constructed, the "cheap" surrogate model is used in place of the "costly" original model to eliminate computational bottlenecks. Until recently, only a small number of studies have applied surrogate modeling techniques to contact problems (Chang et al., 1999; Bouzid and Champliaud, 2003; Lin et al., 2006; Lin et al., 2008). In Chapter 2, we have presented a 2D surrogate-based dynamic contact simulation by using the knowledge of physics. However, such technique was suitable only when contact force is dominated by the translation normal to the contact surface. For 3D knee contact simulations, contact force will be highly sensitive to changes in varus-valgus rotation as well (Fregly et al., 2008). None of the existing techniques is suitable for three-dimensional elastic contact problems where the contacting bodies undergo large relative displacements.

This study proposes a novel surrogate modeling approach for performing computationally-efficient 3D elastic contact analyses within multibody dynamic simulations. The approach uses two key concepts – "sensitive directions" and a "reasonable design space" - to address the unique challenges involved in applying surrogate modeling techniques to elastic contact problems. Sensitive directions define which degrees of freedom in the original model should be sampled using applied loads rather than displacements, while a reasonable design space screens out

infeasible sample points before they are evaluated with the original model. The performance and accuracy of the proposed approach are evaluated by performing thousands of computational wear simulations for a TKR tested in a Stanmore knee simulator machine.

Methods

Surrogate Contact Model Development

Special surrogate model creation techniques are required to apply surrogate modeling methods to elastic contact problems. From a big picture perspective, we treat surrogate contact model creation as a nested process. First, a coarse surrogate model is constructed by evaluating a limited number of sample points with the original contact model. Next, the coarse surrogate model is used to rapidly evaluate the complete set of sample points and screen out infeasible ones. Finally, a fine surrogate contact model is constructed by evaluating all feasible sample points with the original contact model. This approach minimizes the number computationally costly analyses performed by the original contact model during the surrogate model creation process.

The remainder of this section describes our special surrogate model creation process in detail. Eight key assumptions provide the foundation for the subsequent theoretical development:

1. Three-dimensional elastic contact is being modeled between a master body and a slave body possessing general surface geometry and material properties.
2. The pose of the slave body relative to the master body is defined by six pose parameters (i.e., three translations x, y, z along with three rotations α, β, γ using a specified rotation sequence).
3. For each combination of six pose parameters there exists a unique combination of six contact loads (i.e., three forces F_x, F_y, F_z and three torques T_α, T_β, T_γ) experienced by the two contacting bodies in an equal and opposite sense.
4. For any pose, the 6 x 6 matrix defining the sensitivity of the contact loads to pose parameter changes is diagonally dominant (e.g., F_y is most sensitive to changes in y).

5. A deformable body contact model (henceforth called the "original model") is the "computationally costly" contact model to be replaced.

6. A "sample point" is any combination of six pose and load inputs to the original model (e.g., x, F_y, z, α, β, γ) for which the six corresponding load and pose outputs are desired (e.g., F_x, y, F_z, T_α, T_β, T_γ).

7. Given a large number of sample points, repeated static analyses can be performed with the original model to generate input-output relationships to be fitted by the "computationally cheap" surrogate contact model.

8. A surrogate contact model is actually a collection of six separate surrogate models – one for each of the six contact loads.

Based on these foundational assumptions, we have devised a three-step surrogate model development process to address the unique characteristics of elastic contact models.

Step 1) Identify a non-standard sampling method using the concept of sensitive directions. As noted above, the surrogate model creation process involves performing repeated computational experiments with the original contact model to generate input-output pairs to be fitted by the surrogate contact model. The combinations of inputs to be analyzed are selected using design of experiments (DOE), which employs statistical methods to scatter sample points uniformly throughout the bounded six-dimensional design space. Within the context of a multibody dynamic simulation, the desired inputs to a 3D elastic contact model are the six pose parameters and the desired outputs are the six contact loads (Bei and Fregly, 2004). Thus, using a traditional sampling method, the sample points would be combinations of the six pose parameters x, y, z, α, β, γ, with an upper and lower bound placed on each one. The problem with this approach is that physically realistic contact occurs over a thin hyper-surface in six-dimensional pose parameter design space. Consequently, many of the selected sample points will correspond to situations where the opposing surfaces are either out of contact or deeply penetrating. Inclusion of physically unrealistic sample points in the surrogate model fitting process is undesirable since it reduces the accuracy of the resulting surrogate contact model.

To resolve this issue, we propose the concept of "sensitive directions" to modify the traditional sample point definition so as to avoid physically unrealistic situations. When the master and slave body are in contact, some contact loads will be highly sensitive to changes in their corresponding pose parameters, while others will be insensitive. To quantify these sensitivities, we calculate six central difference derivatives and collect the results in a sensitivity vector **s** (Equation 3-1)

$$\mathbf{s} = \left[\partial F_x / \partial x \quad \partial F_y / \partial y \quad \partial F_z / \partial z \quad \partial T_\alpha / \partial \alpha \quad \partial T_\beta / \partial \beta \quad \partial T_\gamma / \partial \gamma \right] \tag{3-1}$$

Due to differences in units, sensitive directions are determined separately for translations (first three entries of **s**) and rotations (last three entries of **s**). If one translational (or rotational) derivative in **s** is significantly larger than the other two, the corresponding direction is deemed a sensitive direction. If two translational (or rotational) derivatives are significantly larger than the third, then two sensitive directions exist. For some situations (e.g., conformal contact between a sphere and a spherical cup of slightly larger radius), all three translational (or rotational) derivatives may be of comparable magnitude, in which case knowledge of the physical situation must be used to determine whether three or zero sensitive directions exist. At least one translational sensitive direction will always exist corresponding to an approximate contact normal direction, and two or more sensitive directions may exist depending on the geometry of the contacting bodies.

Once sensitive directions have been identified, the definition of a sample point is modified such that pose parameters for sensitive directions are replaced by their corresponding contact loads. For example, if *y* is identified as a sensitive direction, then *y* would be replaced with F_y in the sample point definition. With this non-traditional sampling method, sample points become

combinations of pose parameters and contact loads (e.g., x, F_y, z, α, β, γ), with an upper and lower bound placed on each pose parameter and contact load.

Given this modified sample point definition, we generate a specified number of sample points n using a DOE approach. Common approaches include optimal Latin hypercube sampling (Mckay, 1979), Hammersley quasirandom (HQ) sampling (Hammersley, 1960), and face-centered central composite design (Myers and Montgomery, 1995). We choose to use HQ sampling for two reasons. First, HQ sampling provides better uniformity properties than do other sampling techniques for sample points generated within a multi-dimensional hypercube (Kalagnanam and Diwekar, 1997; Diwekar 2003). Second, HQ sample points are generated sequentially rather than simultaneously. Thus, the first m sample points from a larger set of n sample points ($n > m$) will always be approximately equidistant from one another and can be used as a subsample. Furthermore, all existing sample points can be kept if new sample points need to be added.

The repeated static analyses are performed with the original contact model to find the physically realistic poses for the given sample points. During each static analysis, sensitive directions are free to equilibrate under the specified applied loads, while insensitive directions are constrained to the specified pose parameter values. For each static analysis, the resulting loads in sensitive directions are compared to their corresponding applied values. A sample point is chosen only when the maximum percent error of 0.1% is achieved. By the end of static analyses, we ensure that all chosen points free from situations like out of contact or deeply penetrating.

Step 2) Filter out infeasible sample points before evaluation using a coarse surrogate contact model and the concept of a reasonable design space. Though physically unrealistic

sample points are avoided by using the modified sample point definition described above, the resulting sample points may still be outside the envelope of a realistic dynamic simulation (i.e., infeasible sample points). In addition to placing physically realistic upper and lower bounds on the pose parameter and contact load inputs (e.g., x, F_y, z, α, β, γ), we place physically realistic upper and lower bounds on the corresponding contact load and pose parameter outputs (e.g., F_x, y, F_z, T_α, T_β, T_γ). These bounds are estimated by performing a nominal dynamic simulation with the original contact model and expanding the observed range of each pose parameter and contact load by some specified percentage p. Sample points where the corresponding contact model outputs are outside their allowable bounds are deemed infeasible. Similar to the inclusion of unrealistic sample points, inclusion of infeasible sample points in the surrogate model fitting process adversely affects the accuracy of the resulting surrogate contact model. Furthermore, evaluation of infeasible sample points by the original contact model wastes a large amount of CPU time performing computationally costly static analyses whose results will be discarded.

To resolve this issue, we utilize the concept of a "reasonable design space" to screen out infeasible sample points before evaluating them with the original contact model. Generally, the design space can be refined by either directly reducing the range of design variables or by using some computationally inexpensive tool to define irregular design space boundaries (Mack et al., 2007). For the latter case, the reasonable design space approach has been successfully applied to different studies for developing efficient and accurate surrogate-based optimization (Kaufman et al., 1996; Roux *et a*l., 1999; Hosder et al., 2001). To estimate the reasonable design space, we select the first m sample points from the complete set of n sample points generated by HQ sampling. These m sample points are the foundations of reasonable design space approach. First, a static analysis mentioned in previous step is performed with the original contact model for each

of the *m* sample points, and physically unrealistic points are identified and eliminated. Second, six coarse surrogate models are created based on fitting six static analysis outputs (e.g., F_x, y, F_z, T_α, T_β, T_γ) of the remaining points (< *m*) as functions static analysis inputs (e.g., x, F_y, z, α, β, γ). Third, six coarse surrogate models are evaluated repeatedly to determine the feasibility of the *n-m* sample points not used to construct it. Fourth, the final set of all feasible sample points (<< *n*) is used to construct a fine surrogate contact model. Since a large number of the *n-m* sample points will turn out to be infeasible, screening these points with the coarse surrogate contact model provides significant computational savings.

Step 3) Construct a fine surrogate contact model using the feasible subset of the original sample points. Fitting a fine surrogate contact model to the feasible sample points is complicated by the fact that contact model outputs and inputs are different for fitting versus sampling. To illustrate this issue, consider an elastic contact model whose only sensitive direction is *y* translation. During sampling of the original contact model, outputs would be F_x, y, F_z, T_α, T_β, T_γ while inputs would be x, F_y, z, α, β, γ. During fitting of the surrogate contact model, outputs would be F_x, F_y, F_z, T_α, T_β, T_γ and inputs would be x, y, z, α, β, γ, as required by the dynamic simulation. Altered interpretation of outputs and inputs between sampling and fitting is not problematic, since at the end of the sampling process, we are left with a table containing six pose parameters and six contact loads for each sample point.

With this issue in mind, the final surrogate model creation process proceeds as follows. Repeated static analyses are performed with the original contact model for all feasible sample points not yet evaluated. Since the coarse surrogate models may have wrongly identified some infeasible sample points as feasible, sample points whose outputs are outside their allowable bounds are discarded. The final set of all feasible sample points (<< *n*) is used to construct the

final surrogate contact model. Contact loads corresponding to sensitive directions are fitted as a function of the six pose parameters (e.g., $F_y = f(x, y, z, \alpha, \beta, \gamma)$), while contact loads corresponding to insensitive directions are fitted as a function of contact loads in sensitive directions and pose parameters in insensitive ones (e.g., $F_x = f(x, F_y, z, \alpha, \beta, \gamma)$). This approach keeps the fitting process as close to the sampling process as possible without solving a nonlinear root-finding problem for the sensitive directions (e.g., given $y = f(x, F_y, z, \alpha, \beta, \gamma)$, find F_y such that $f(x, F_y, z, \alpha, \beta, \gamma) - y_c = 0$, where y_c is the current value of y in the dynamic simulation). Actual fitting of the six input-output relationships is performed using Kriging (Krige, 1951), as our preliminary studies revealed that Kriging produces more accurate dynamic simulation results than do a variety of other surrogate model fitting methods (e.g., polynomial response surface and support vector regression). Once a Kriging-based surrogate contact model has been constructed, it is used to calculate six contact loads at each time instant during a dynamic simulation given the six pose parameters for the contacting bodies.

Surrogate Contact Model Evaluation

The proposed surrogate modeling approach was evaluated by performing Monte Carlo analyses and forward dynamic simulations using a multibody dynamic model of a cruciate-retaining commercial knee implant (Depuy Orthopedics, Warsaw, IN) mounted in a Stanmore knee simulator machine (Walker et al., 1997). The model possessed six degrees of freedom (DOFs) relative to ground: tibial anterior-posterior translation x, tibial medial-lateral translation z, tibial internal-external rotation β, femoral superior-inferior translation y, femoral varus-valgus rotation α, and femoral flexion-extension γ (Figure 3-1). Similar to an actual Stanmore machine, γ was motion controlled, x, y, and β were load controlled, and z and α were left free, where all controlled quantities used ISO standard input curves (DesJardins et al. 2000).

A 3D elastic foundation (EF) contact model (An et al., 1990; Blankevoort et al., 1991; Li et al., 1997; Pandy et al., 1998; Fregly et al., 2003) of the knee implant was used to calculate contact loads between the femoral component and tibial insert given the six pose parameters for the contacting bodies. The EF model utilized linear material properties (Young's modulus = 463 MPa, Poisson's ratio = 0.46; Bei and Fregly, 2004) and surface geometry taken from manufacturer computer-aided design models, with contact occurring on the medial and lateral sides of the tibial insert. Symbolic dynamics equations for the system were derived using Kane's method (Kane and Levinson, 1985) and Autolev (OnLine Dynamics, Sunnyvale, CA). These equations and the EF contact model were incorporated into a Matlab program (The Mathworks, Natick, MA) that was used to perform forward dynamic simulations with Matlab's stiff numerical integrator ode15s.

The surrogate modeling approach presented is a general algorithm, and it can be applied to any number of elastic contacts. To demonstrate this ability during the surrogate model evaluation, two respective surrogate contact models were constructed to replace the medial and lateral contacts occurring on the tibial insert. These two models were created using three-step process outlined above.

Step 1) Identify a non-standard sampling method using the concept of sensitive directions. With the implant components in a nominal anatomic pose and both sides barely touching, we calculated the sensitivity vector **s** defined in Equation 3-1 and identified two sensitive directions. Specifically, contact force F_y and contact torque T_α were found to be highly sensitive to small variations in superior-inferior translation y and varus-valgus rotation α, respectively. Thus, sample points were defined as $x,\ F_y, z,\ T_\alpha, \beta, \gamma$. Upper and lower bounds for the sample points were determined by performing 16 dynamic contact simulations with the EF

contact model using motion and load input curves at the extremes of their allowable variations for the subsequent Monte Carlo analyses. After expanding the resulting bounds by p = 50%, we generated n = 2000 sample points using the HQ sampling method (Figure 3-2A), where the values for p and n were chosen based on previous experience with surrogate contact models (Lin et al., 2008).

Step 2) Filter out infeasible sample points before evaluation using the concept of a reasonable design space. To estimate which sample points were within the reasonable design space, we selected m = 500 sample points from the complete set of 2000. We then performed 500 static analyses with the EF contact model to calculate the corresponding outputs F_x, y, F_z, α, T_β, T_γ, where each static analysis required approximately 24 seconds of CPU time on a 3 GHz Pentium IV PC (Figure 3-2B). 18 sample points were identified as unrealistic, leaving 482 sample points for constructing coarse surrogate models by fitting six outputs (i.e., F_x, y, F_z, α, T_β and T_γ) as functions of sample inputs (Figure 3-2C). After performing a cross validation analysis with these 482 sample points, we selected a cubic polynomial and Gaussian correlation function to construct all surrogate models using the DACE Toolbox for Matlab (Lophaven et al., 2002). Since each sample points contained information of medial and lateral side contact loads, two sets of six coarse surrogate models were created to predict their corresponding values for the remaining 1500 sample points (Figure 3-2D). 934 (937) of these points were found to be infeasible for medial (lateral) contact, leaving 566 (563) additional sample points to be evaluated with the EF contact model (Figure 3-2E).

Step 3) Construct the final surrogate contact model using the feasible subset of the original sample points. Considering the overlap between three groups of sample points (566, 563, and first 500 sample points), only 560 additional static analyses were performed with the EF

contact model for the feasible sample points identified in the previous step (Figure 3-2F). The CPU time required for 1060 (=500+560) static analyses was 6 hours. The results of these static analyses showed that 500 (560) out of 566 (563) sample points were feasible for creating the medial (lateral) part of the final surrogate contact model. Each part fitted six contact loads as a function of the following inputs (Equation 3-2, Figure 3-2G):

$$\begin{aligned} F_x &= f(x, F_y, z, T_\alpha, \beta, \gamma) \\ F_y &= f(x, y, z, \alpha, \beta, \gamma) \\ F_z &= f(x, F_y, z, T_\alpha, \beta, \gamma) \\ T_\alpha &= f(x, y, z, \alpha, \beta, \gamma) \\ T_\beta &= f(x, F_y, z, T_\alpha, \beta, \gamma) \\ T_\gamma &= f(x, F_y, z, T_\alpha, \beta, \gamma) \end{aligned} \tag{3-2}$$

Since x, y, z, α, β, γ are the inputs, Equation 3-2 indicates that contact loads in the sensitive directions (i.e., F_y and T_α) must be calculated before contact loads in the insensitive directions. The final surrogate contact model was composed of two sets of these six Kriging models and was incorporated into the multibody dynamic model in place of the EF contact model (Figure 3-2H). During the dynamic simulation, the total contact loads are generated by summarizing Kriging-based medial and lateral contact loads.

Six additional Kriging models were created to fit medial and lateral center of pressure (CoP) location (3 components per side) as functions of the sample inputs. The CoP predictions and previously calculated contact forces were used to calculate medial and lateral wear volume. Details of the process are given by Lin et al (2008). Briefly, Archard's wear law (Equation 2-4) with a wear factor of 1×10^{-7} mm^3/Nm (Fisher et al., 1994) was used to predict wear volume on each side from the time history of contact force and CoP slip velocity along with the time increment for numerical integration. Once the CoP location was predicted for a given pose, the first time derivative of the pose parameters and the kinematic concept of two points fixed on a

rigid body were used to calculate the corresponding CoP slip velocity (Kane and Levinson, 1985).

To evaluate the performance of surrogate-based dynamic simulation, we performed five Monte Carlo analyses to investigate how realistic variations in motion and load inputs affect predicted wear volume. The first four Monte Carlo analyses varied the three load inputs and one motion input separately while the fifth analysis varied all input profiles together. Principal component analysis was used to create realistic variations of the input curves based on experimentally observed variations for eight different implant designs tested in a Stanmore simulator machine (DesJardins et al. 2000; Lin et al., 2008). Each new input curve was generated by selecting weights between 0 and 1 for the first two principal components, which captured 98% of the variability in each type of curve. Each Monte Carlo analysis was performed on a 3 GHz Pentium IV PC and involved at least 1000 forward dynamic simulations incorporating the surrogate contact model. The convergence criterion for each analysis was met when the mean and coefficient of variation (i.e., 100*standard deviation/mean) for the last 10% of the wear predictions were within 2% of the final mean and coefficient of variation (Fishman 1996; Valero-Cuevas et al., 2003).

To evaluate surrogate contact model accuracy, we performed eleven dynamic wear simulations with both the surrogate contact model and the EF contact model used to create it. A nominal dynamic simulation was first performed using the ISO standard input curves and the simulation results from two models compared quantitatively. We then performed ten additional EF-based and surrogate-based dynamic simulations to evaluate the extremes in input curve variations used during the Monte Carlo analyses. Specifically, for each Monte Carlo analysis, we performed two dynamic simulations with each contact model where each input curve was

generated using either the maximum or minimum sum of the component weights. The ten pairs of results were compared quantitatively by calculating root mean square errors (RMSE) and maximum absolute errors (MAE) for predicted pose parameters, medial and lateral contact loads, and wear volumes. The CoP-based approach for calculating wear volume was also evaluated by calculating wear volume with the EF model using both an element based approach (Fregly et al., 2005; Zhao et al., 2008) and the CoP-based approach (Lin et al., 2008).

Results

For the Monte Carlo analyses, the results revealed small variations in input motion relative to the ISO standard had different effects on predicted wear volume (Figure 3-3). The predicted wear volume was more sensitive to changes in flexion-extension motion than to changes in input loads. The most variation was obtained when all of the input curves were varied simultaneously. Each Monte Carlo analysis performed with the surrogate contact model required 1.4 hours of CPU time compared to an estimated 284 hours with the EF contact model. As for a single dynamic simulation, the CPU time required for EF and surrogate contact model was 17 minutes and 5 seconds, respectively.

The comparison between multiple EF-based and surrogate-based dynamic simulations demonstrated that the surrogate contact model accurately reproduced the contact kinematics, contact loads, and wear volumes predicted with the EF contact model. The nominal dynamic simulation results showed similar behavior between the surrogate and EF contact model (Figure 3-4 to 3-6). Similar to the nominal dynamic simulation, the results of the other ten surrogate-based dynamic simulations were close to the EF-based simulation results. On average, the RMSE and MAE for 3D translations/rotations were less than 0.2 mm/0.6 deg and 0.3 mm/ 1.3 deg, respectively. For the medial contact forces/torques, RMSE and MAE were less than 20 N/0.540 N-m and 46 N/1.11 N-m, respectively. For the lateral contact forces/torques, RMSE and

MAE were less than 20 N/0.561 N-m and 41 N/1.12 N-m, respectively. For both the surrogate contact model and the EF contact model, the CoP-based approach for predicting wear volumes reproduced the element-based EF results to within 2% error (Table 3-1).

Discussion

Computational speed is a major limiting factor for performing sensitivity and optimization studies of dynamic contact problems. This paper has presented a novel surrogate modeling approach for performing efficient dynamic contact simulations. The approach requires use of the contact surfaces and the knowledge of sensitive directions to generate the sample points while no computationally expensive geometry evaluation is required during the dynamic simulation. The methodology was successfully applied to replace the EF-based artificial knee contact model created from the manufactured CAD data. The evaluation demonstrated the computational efficiency of the approach by performing 5 Monte Carlo analyses and highlighted that the wear predicted form a simulator machine was sensitive to the changes in machine inputs. The results of 11 dynamic simulation comparisons demonstrated that the presented surrogate contact model accurately reproduced the EF-based dynamic simulation results even when the input curves were varied from their standard values during the Monte Carlo analyses.

The proposed surrogate contact model is based on two key concepts. First, knowledge of sensitive directions is used to generate preliminary sample points without producing physically unrealistic outputs. This key concept is based on the assumption that the value of each contact load is dominated by the value of the corresponding pose parameter. This assumption implicitly assumed that the derivatives of contact forces/torques with respect to the non-corresponding translations/rotations were not significant. To validate this assumption, the derivatives with respect to the non-corresponding pose parameters were calculated for the proposed TKR design and the results proved that F_y and T_x were truly dominated by the values of Y and γ, respectively.

The results were also consistent with previous study done on the same TKR design (Fregly et al., 2008) reported that the variations in superior-inferior translation and varus-valgus rotation have much more influence on the contact quantities than did variations in the remaining DOFs. Although this assumption is valid for this study, there are other cases in which it fails. For example, a hip joint (a sphere in a spherical cup) would have three sensitive translation directions and each sensitive direction might have less influence on the corresponding force than other two forces. Under this circumstance, we may no longer sample the corresponding force but instead sample one of the other two forces. Future data collected from a hip joint should be used to investigate this issue further.

Second, a reasonable design space approach is developed to refine the design space and generate an accurate surrogate contact model. It screened out infeasible points from the design space and provided more accurate surrogate models. This study used the initial surrogate models developed from a subset of the sample points to screen all sample points before they were used in the computational experiments. These initial surrogate models may not provide highly accurate predictions, but they efficiently identified most of the appropriate sample points even based on their limited prediction abilities. About 88% (i.e., 500 out of 563 and 496 out of 566) of the accepted sample points turned out to be within the reasonable design space. The additional static analyses performed on the rest of the 2000 sample points further indicated that only 5% of the "bad" points (including the unrealistic and infeasible points) were not captured by the initial surrogate models.

A hybrid approach was utilized to develop the final surrogate contact model for couple reasons. Theoretically, two fitting approaches could be used to develop the final surrogate contact model: sample-based fitting and use-based fitting. For sample-based fitting, we fitted F_x, F_z, T_y,

T_z, and two sensitive directions (i.e., Y and α) as functions of sample inputs and used a root-finding technique to solve for F_y and T_x during the forward dynamic simulation. However, the extra computational time spending on the root-finding made this approach less appealing in practice. For use-based fitting, we fitted all contact loads as functions of six pose parameters and then performed the forward dynamic simulation by manipulating all six pose parameters simultaneously. Although the preliminary results showed that this approach successfully completed a dynamic simulation within 5 seconds of CPU time, the accuracy was not as good as the current approach. Thus, given the weakness of both fitting approaches, we developed a hybrid approach to combine these two approaches for performing fast and accurate dynamic simulation.

Use of the surrogate contact model to replace repeated contact analyses is worthwhile for several reasons. The primary benefit is improved computational efficiency. Although the computational cost of performing repeated static analyses is inevitable, this cost is pay up front cost since the sampling process is only performed once. In addition, such cost would be recovered quickly when performing sensitivity or optimization studies. In the presented application, the CPU time spent on 1060 static analyses was approximately 6 hours. However, the use of the surrogate contact model in place of the EF model reduced Monte Carlo analysis time from 284 hours to 1.4 hours, a nearly 99% reduction. While the application has demonstrated that surrogate contact models can be beneficial for sensitivity studies, the greatest computational benefit is likely to occur for optimization studies. For stochastic sensitivity studies, the mean value method (Wu et al., 1990) has been used to reproduce the Monte Carlo wear predictions while requiring only 6% of computational time (Pal et al., 2008). In contrast, no good methods exist for reducing the number of repeated simulations in optimization studies. For those situations, the factor of 200 reduction in computation time (i.e., 17 minutes to 5 seconds) could mean the difference between

an impossible and an achievable optimization study. Other benefits of using surrogate contact models were provided in Chapter 2.

The sensitivity analysis results demonstrate the importance of controlling machine inputs closely, especially the femoral flexion angle. It should be noted that the current surrogate-based wear prediction did not include creep information. However, the absence of creep calculation should not affect the wear volume prediction significantly (Zhao et al., 2008). The Monte Carlo analyses revealed when small variations were imposed on the input motions and loads, only variations in the femoral flexion angle resulted in large changes in predicted wear volume (Figure 3-5). This finding was consistent with results from our previous study where the similar Monte Carlo analyses were applied to a similar TKR design constrained to planar motion in the Stanmore machine (Lin et al., 2008). Pal et al. (2008) also discovered a similar result where they developed a probabilistic wear prediction model to identify that femoral flexion–extension alignment was one of the parameters most affecting predicted wear.

The current modeling approach is limited in at least three aspects. First, the results of current central finite difference scheme were influenced by the coordinate system definition of the master body. For instance, X translation might become as sensitive as Y translation once we rotated our tibial insert coordinate system 45 degree about its z-axis. However, there is a redundancy between these two sensitive directions (i.e., X and Y translation) since they both are highly related to contact force F_y in original y-axis direction. To minimize the redundancy between the sensitive directions, the technique likes principal component analysis could be used to determine a "principal sensitive direction" from any highly correlated sensitive directions. More extensive data would be required to draw any conclusions about the effectiveness of principal component analysis for reducing the redundancy.

Second, new surrogate contact models must be generated if geometries or material properties of the contact bodies are changed. However, this limitation is not serious for the presented application since recent studies have reported that predicted wear volume (but not wear area or depth) is relatively insensitive to whether or not the surface geometry is changed gradually (Knight et al., 2007; Zhao et al., 2008). Furthermore, the additional design variables process (e.g., make material property an additional surrogate model input) could be added into the surrogate modeling process.

Third, dynamic simulations obtained from the computationally expensive model must be performed to evaluate the accuracy of the surrogate contact model. This limitation prevents us from evaluating all surrogate-based dynamic simulations performed in Monte Carlo analysis. In general, the predictive capability of a surrogate model is evaluated either by using the "non-training points" (i.e., points not used to develop a surrogate model) or by performing the prediction error sum of squares analysis (Martens and Naes, 1989) on the training points. However, these two methods may be insufficient for dynamic contact problems since it is difficult to draw firm conclusions about the contributions that the predictive capability of a surrogate model make to accurately reproduce the original dynamic simulation. One of the possible alternatives to evaluate the surrogate-based dynamic simulation without relying on the computationally expensive model is to perform a probability analysis. Such analysis would establish likelihood measures of a new pose parameter given by the forward dynamic simulation being consistent with the training points. Higher probability indicates greater confidence that the surrogate model will produce a more accurate prediction, and consequently improve the quality of dynamic simulation results.

In summary, this study has presented a detailed surrogate contact modeling approach for significantly improving the computational speed of 3D dynamic contact simulations. The evaluation demonstrated the advantage of using the surrogate contact model to perform sensitivity and optimization studies incorporating multiple deformable contacts. In the biomechanics field, such computationally efficient model would be ideal for the simultaneous predictions of contact forces and muscle forces since optimization requiring thousands of contact simulations is the essential part of the prediction process. Extension of the current approach to a full-leg musculoskeletal model incorporating multiple surrogate contact models is one of the ongoing research topics. Once surrogate contact models are developed for all major lower extremity joints, it should be possible to perform dynamic simulations in a short amount of CPU time.

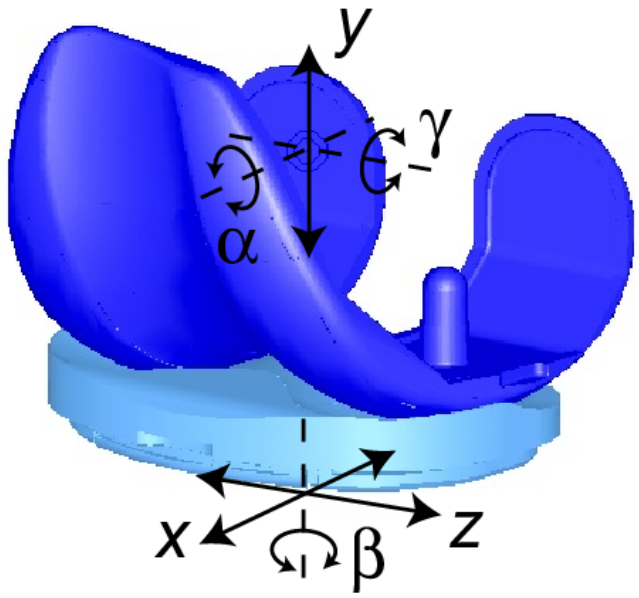

Figure 3-1. A 6 degree of freedom total knee replacement contact model. The tibial insert (light blue object) poses three DOFs (anterior-posterior translation x, medial-lateral translation z, internal-external rotation β) and the femoral component (dark blue object) poses another three DOFs (superior-inferior translation y, varus-valgus rotation α, and flexion-extension γ).

Table 3-1. Comparison between the wear volume predictions (mm^3) made with the elastic foundation contact model and the surrogate contact model for dynamic simulations with different input curves.

Quantity varied	Sum of component weights	Elastic foundation contact model	Surrogate contact model	Difference (%)
F_x		28.44	28.51	0.24%
F_y		30.31	30.27	0.14%
T_y	Maximum	18.14	18.24	0.56%
Flex		17.10	17.23	0.72%
All		29.76	29.52	0.80%
F_x		29.49	29.20	1.02%
F_y		30.19	29.73	1.53%
T_y	Minimum	27.63	27.37	0.96%
Flex		27.08	26.90	0.66%
All		30.11	30.08	0.10%

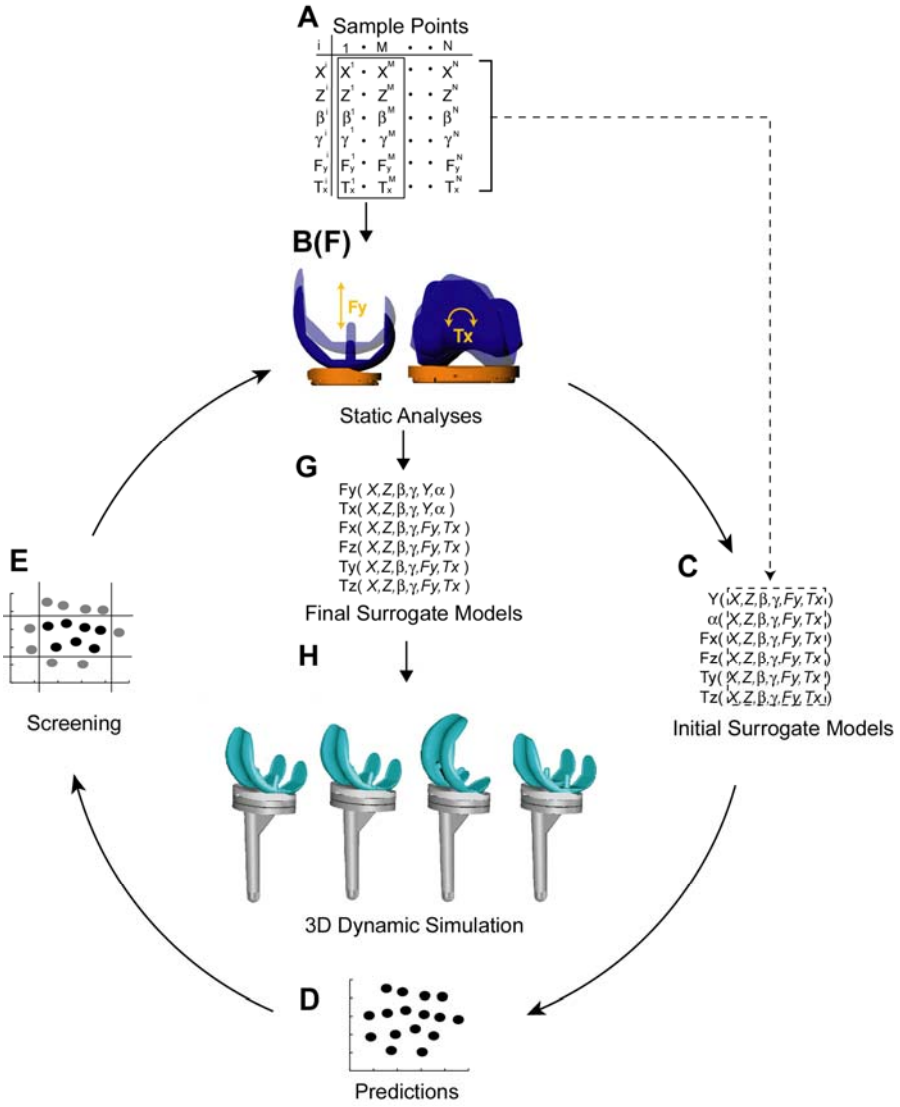

Figure 3-2. Overview of the surrogate modeling approach. A) N points are sampled in the (X, Z, β, γ, F_y, T_x) design space where the first M points are used to generate the initial surrogate models. B) Static analysis is performed at the selected sample points using an elastic foundation (EF) contact model. C) The initial surrogate models are developed based on the first M sample points. D) The initial surrogate models are used to predict the static analyses results for N sample points. E) The predictions are screened based on the ranges defined by the EF-based dynamic simulation results. F) Additional static analyses are performed to obtain the outputs of interest for all remaining sample points. G) The final surrogate models are developed based on the remaining sample points. H) Given the current values of six pose parameters, the surrogate contact model calculates the contact loads throughout the dynamic simulation.

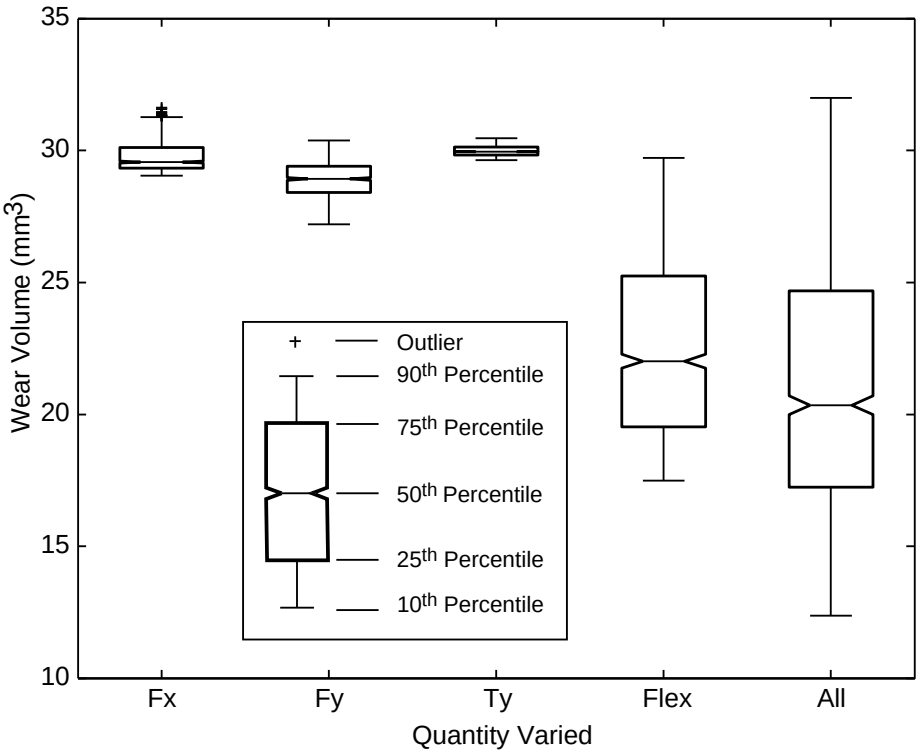

Figure 3-3. Box plot distribution of predicted wear volume generated by four Monte Carlo analyses using the surrogate contact model.

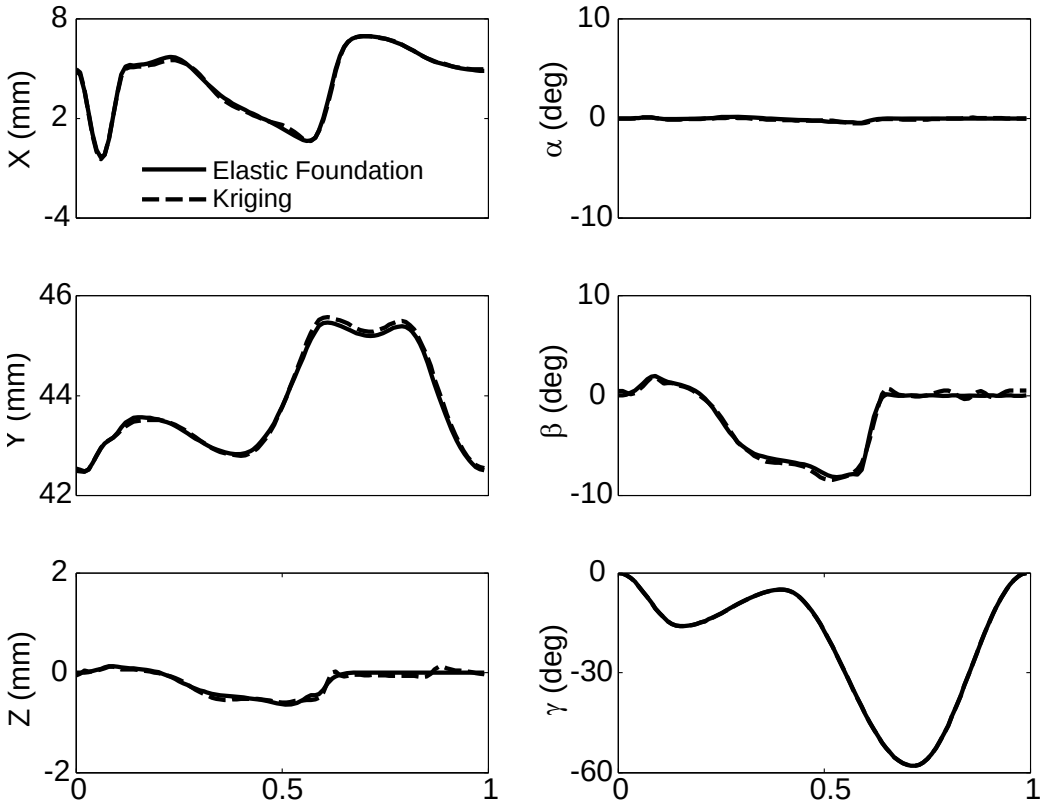

Figure 3-4. Comparison between the six pose parameters made with the elastic foundation contact model and the surrogate contact model for the nominal dynamic simulation.

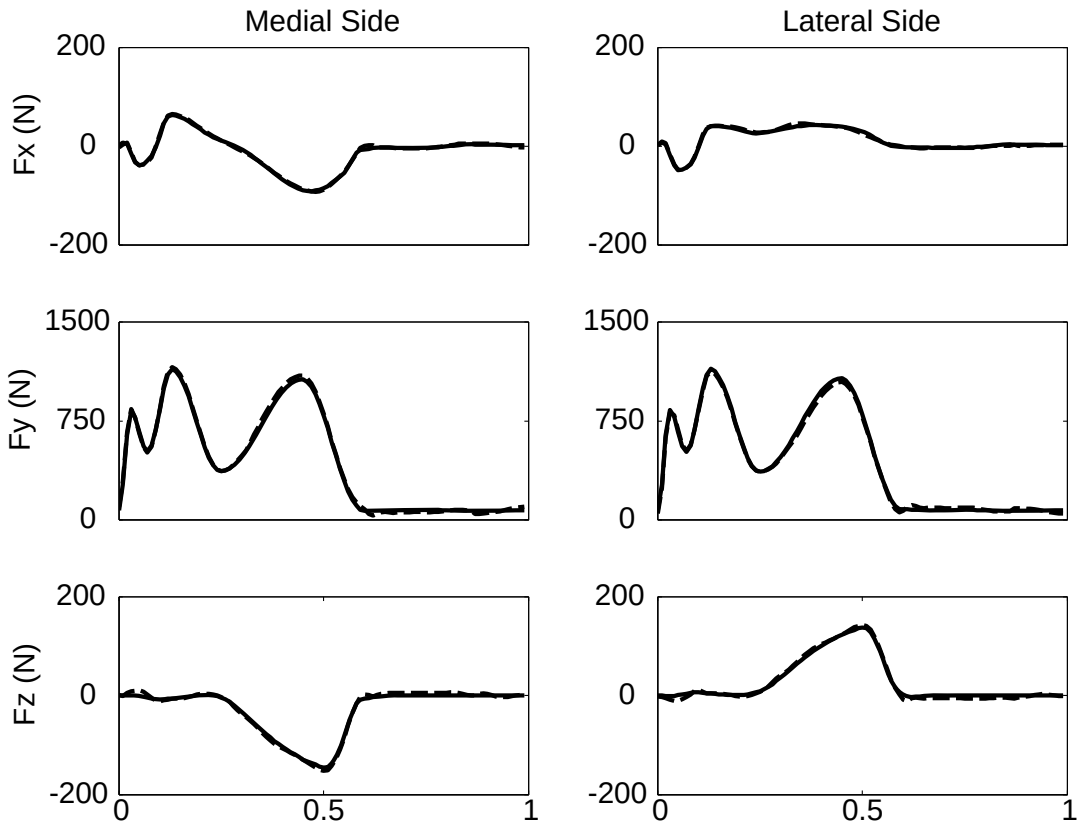

Figure 3-5. Comparison between the medial-lateral contact forces made with the elastic foundation contact model and the surrogate contact model for the nominal dynamic simulation.

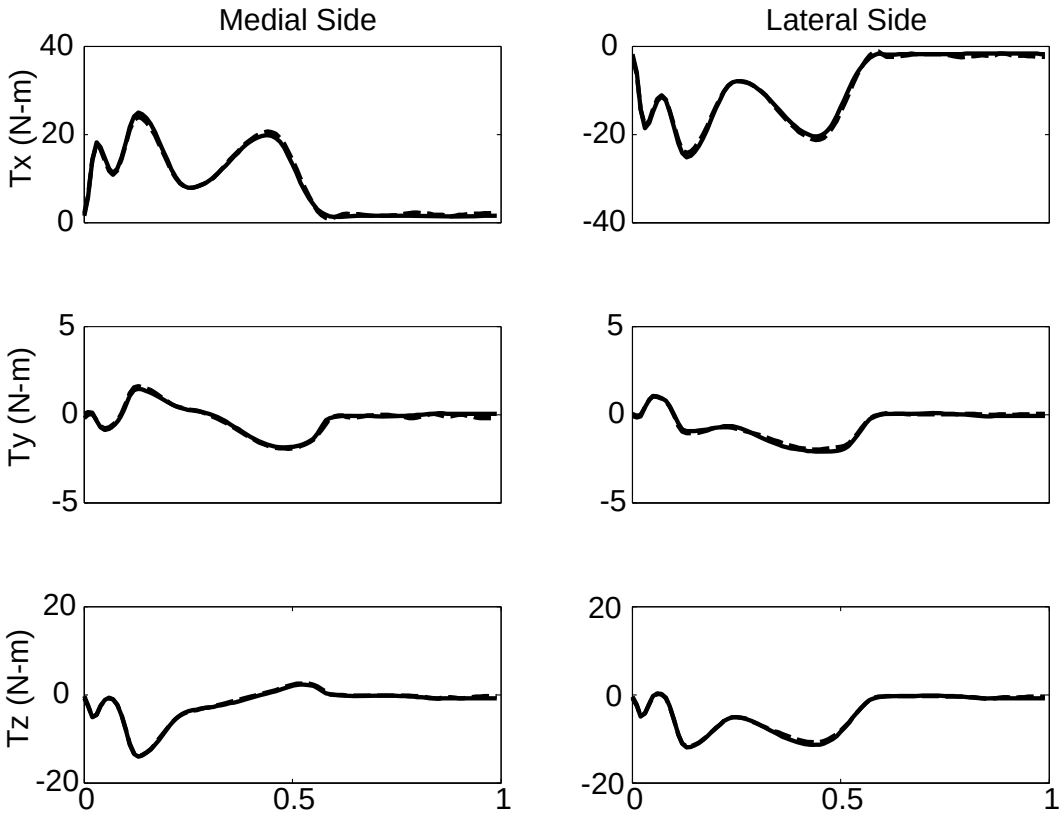

Figure 3-6. Comparison between the medial-lateral contact torques made with the elastic foundation contact model and the surrogate contact model for the nominal dynamic simulation.

CHAPTER 4
SIMULTANEOUS PREDICTION OF MUSCLE AND CONTACT FORCES IN THE KNEE DURING GAIT

Introduction

Knowledge of muscle forces during walking is necessary to characterize muscle coordination and function, which are factors in neurological disorders such as cerebral palsy and stroke (Schmidt et al., 1999; Higginson et al., 2006; Gibson et al., 2007; Sujith, 2008), and to characterize joint and soft-tissue loading, which are factors in bone and joint diseases such as osteoarthritis and patellofemoral pain syndrome (Hurwitz et al., 2000; Fregly et al., 2007; Milner et al., 2007). For all of these disorders, the ability to determine muscle and joint loads accurately in vivo could lead to significant improvements in clinical outcome. Unfortunately, direct measurement of muscle forces in vivo is possible only under limited special circumstances (Schuind et al., 1992; Dennerlein et al., 1998; Finni et al., 1998; Ishikawa et al., 2005), and similarly for joint contact loads (Stansfied et al., 2003; D'Lima et al., 2006). Due to these limitations, musculoskeletal computer models have become the primary scientific approach for predicting muscle forces and contact loads during human movement (Anderson and Pandy, 2001a; Buchanan et al., 2004; Higginson et al., 2006; Erdemir et al., 2007).

A primary challenge to developing such predictions is the non-uniqueness of the calculated muscle forces, often referred to as the "muscle redundancy problem." Since more muscles act on the skeleton than the number of degrees of freedom in the skeleton, an infinite number of possible muscle force solutions exist. Consequently, optimization is commonly used to solve the indeterminate problem (Kaufman et al., 1991; Anderson and Pandy, 2001; Erdemir et al., 2007). This approach starts by calculating the net reaction loads (i.e., three forces and three torques) at each joint using a dynamic musculoskeletal model and movement data collected from a patient. A common constrained optimization problem is then used to estimate muscle forces by

minimizing the sum of squares of muscle activations with constraints that muscles balance some (but not all) of the net reaction loads. In physiological terms, the net loads are produced by a combination of muscle and contact forces. Thus, if a contact model is not included explicitly in the musculoskeletal model, then assumptions need to be made about which loads are generated primarily by muscles (i.e., little contribution from contact loads), and these loads are the only ones that can be used as constraints. In contrast, if a contact model is included explicitly in the musculoskeletal model, then all six net loads can theoretically be used as constraints in the optimization problem formulation, thereby eliminating the need for assumptions about the lack of contact contributions to particular loads.

To date, no study has included a contact model explicitly in the musculoskeletal model to predict knee joint muscle forces and contact loads simultaneously during the optimization process. Previous musculoskeletal models of the knee generally used constraint-based joints (e.g., pin joint, ball-and-socket joint) with limited degrees of freedom (DOFs) when in reality the knee joint possesses 12 DOFs (6 for the tibiofemoral joint and 6 for the patellofemoral joint). Despite this limitation, some studies have reported that increasing the DOFs in a constraint-based knee model from 1 to 3 significantly affects the predicted muscle forces (Buchanan et al., 1996; Glitsch et al., 1997; Li et al., 1998). Though this issue can be eliminated by replacing each constraint-based joint with a geometry-based joint (i.e., a contact model), the computational cost of solving repeated contact problems currently makes such an approach infeasible (Fregly et al., 2005).

This chapter compares simultaneous muscle and contact force predictions in the knee generated using different optimization problem formulations with and without matching of in vivo medial and lateral contact force measurements as constraints. The goals of this chapter are

two-fold. The first is to evaluate the ability of a musculoskeletal model incorporating surrogate contact models to reproduce optimization predictions of muscle and contact forces generated using the same musculoskeletal model but incorporating elastic foundation contact models. The second is to utilize the musculoskeletal model with surrogate contact to evaluate the ability of eight optimization problem formulations to reproduce in vivo contact force measurements provided by an instrumented knee implant. A 12 DOF surrogate contact model of the tibiofemoral (TF) and patellofemoral (PF) joints is created and incorporated into the musculoskeletal model for performing the optimizations. The predicted muscle forces are evaluated using in vivo medial and lateral contact force measurements from a single patient with an instrumented knee implant (D'Lima et al., 2006; D'Lima et al., 2007; Zhao et al., 2007b).

Methods

Data Collection

Experimental gait data were collected from a single patient (male, age 80 years; mass 68 kg) implanted with an instrumented total knee replacement (TKR) in his right leg. Institutional review board approval and patient informed consent were obtained. The instrumented knee implant consisted of four uniaxial force transducers, a micro transmitter, and an antenna ((D'Lima et al., 2005). In vivo tibial force data were recorded simultaneously with video motion (Motion Analysis Corporation, Santa Rosa, CA) and ground reaction data when the patient performed different walking trials. For each trial, the distribution of axial contact forces between the medial and lateral compartments of the knee was calculated from the four force transducer measurements using validated regression equations (Zhao et al., 2007a). A modified Cleveland Clinic marker set with additional markers placed on the feet was used for the video motion data collection. Static markers placed over the medial and lateral epicondyles of the knee and the

medial and lateral malleoli of the ankle were used along with dynamic markers to define segment coordinate systems and provide three-dimensional (3D) movement data.

Inverse Dynamics Loads

Net forces and torques (henceforth referred to as inverse dynamics loads) at the knee joint were calculated using a published bottom-up inverse dynamics approach (Reinbolt et al., 2008). In brief, a four-segment leg model was developed specifically for the patient based on the collected surface marker data. The four segments consisted of the foot, shank, thigh, and pelvis. The foot was treated as the root segment connected to the laboratory-fixed coordinate system via a six DOF joint. Two non-intersecting pin joints were used to model the ankle (van den Bogert et al., 1994), a single pin joint was used to model the knee (Reinbolt et al., 2005),and a ball-and-socket joint was used to model the hip. A series of optimizations were applied to the model to calibrate its inertia parameter values (body segment masses, mass centers, and central principal moments of inertia) and joint parameter values (positions and orientations of ankle, knee, and hip joint axes in the body segments) by minimizing the errors between model and experimental marker locations (Reinbolt et al., 2005). The optimized tibiofemoral (TF) joint center was defined as the point on the optimized knee axis at the midpoint between the medial and lateral femoral epicondyle markers. Six inverse dynamics loads at the knee for one cycle (right heel strike to subsequent right heel strike) of a selected gait trial were then calculated about the optimized TF joint center and expressed in the shank reference frame. Corresponding medial, lateral, and total tibial contact forces for the same gait trial were provided by the instrumented knee implant.

Musculoskeletal Model

A composite implant/bone/muscle knee model was constructed based on medical imaging data obtained from two subjects. The model consisted of computer-aided design (CAD) models

of the knee implant components, geometric models of four bones (femur, patella, tibia, and fibula), and origin and insertion locations for eleven muscles and the patellar ligament (Figure 4-1). Coarse bone models were constructed for the patient using his post-surgery computed tomography (CT) data. Both the bones and the metallic implant components were segmented from the CT images, allowing the implant CAD models to be aligned to the segmented outlines of the coarse bone models. Fine bone models with corresponding muscle and ligament attachment points were constructed for a healthy subject of similar stature using magnetic resonance imaging (MRI) data. The healthy subject's fine bone models were aligned to the segmented outlines the patient's coarse bone models, resulting in a composite geometric model consisting of the implant CAD models and the fine bone models with corresponding muscle and ligament attachment points. All alignment tasks were performed to minimize distances between aligned objects using the commercial reverse engineering software Geomagic Studio (Raindrop Geomagic, Research Triangle Park, NC). Once the registration process was finished, the kinematic structure of the inverse dynamics leg model was surperimposed onto the composite geometric model by aligning the ankle, knee, and hip centers of the two models.

Eleven major muscles spanning the knee were included in the model: Vastus medialis (VM), vastus lateralis (VL), vastus intermedius (VI), rectus femoris (RF), semimembranosus (SM), semitendinosus (ST), biceps femoris long head (BFLH), bicpes femoris short head (BFSH), tensor fascia latae (TFL), gastrocnemius medial head (GM), and gastrocnemius lateral head (GL). Each muscle was activated independently except for medial hamstrings (group of SM and ST) and vasti (group of VM, VI, and VL), which were each controlled by a single muscle activation, with muscle activation being used as an indicator of neural control input. Muscle force was generated by multiplying each muscle activation by its corresponding peak isometric

force, as a previous study has shown that this simple muscle model produces nearly identical results to those of more complex muscle models when gait is being simulated (Anderson et al., 2001b). The patellar ligament was modeled as three parallel springs with a total stiffness of 2000 N/mm (Reeves et al., 2003).

The TF and patellofemoral (PF) joints were each modeled as 6 DOF joints with elastic contact (Figure 4-2). The femoral component was fixed to ground and the tibial insert was connected to the femoral component via a 6 DOF joint (generalized coordinates $X_T, Y_T, Z_T, \alpha_T, \beta_T, \text{and } \gamma_T$ for the three translations and rotations, respectively). The patella was also connected to the femoral component via a 6 DOF joint (generalized coordinates $X_P, Y_P, Z_P, \alpha_P, \beta_P, \text{and } \gamma_P$). These two 6 DOF joints were used to quantify the relative position and orientation of each pair of articulating implant components for contact calculations. For the TF joint, anterior-posterior translation, internal-external rotation, and flexion-extension were prescribed to match the kinematics measured from video motion, since contact force calculations are insensitive to small errors in these DOFs (Fregly et al., 2008). In contrast, for the PF joint, none of the 6 generalized coordinates was prescribed. No relative motion was allowed between the thigh and femoral component or between the shank and tibial insert. For each time frame, the values of the 9 free generalized coordinates were determined based on static equilibrium of the shank and patella produced by the applied inverse dynamics loads, contact loads, muscle forces, and patellar ligament force.

Surrogate Contact Model

Two surrogate contact models, one for the TF joint and one for the PF joint, were developed to perform repeated contact analyses efficiently. For each joint, a surrogate contact model was created using the corresponding implant CAD geometry, a validated elastic

foundation (EF) contact model (Fregly et al., 2003; Bei et al., 2004; Zhao et al., 2008), and the methods presented in Chapter 3. In brief, we used a three-dimensional EF contact model with nonlinear material properties (Fregly et al., 2007) to generate sample points for creating TF and PF surrogate contact models. For both contact models, superior-inferior translation and varus-valgus rotation were identified as sensitive directions since the contact loads were highly sensitive to small variations in these generalized coordinate values. We therefore sampled force and torque (F_y and T_x) in these two sensitive directions and pose (X, Z, β, and γ) in the remaining four insensitive directions. All sampling was again performed using a Hammersley quasirandom sequence. Physically realistic upper and lower bounds on the sample point inputs (i.e., X, Z, β, γ, F_y and T_x) were determined by performing a sequence of static analyses across the gait cycle using a knee model incorporating EF contact models. A simplified version of model was created for this purpose where vasti was the only muscle and possessed a constant force of 200 N. For each time frame, the pose of the TF joint was prescribed to reproduce the in vivo medial and lateral tibial contact force measurements while the pose of the PF joint was determined based on static equilibrium of the patella.

A reasonable design space approach as described in Chapter 3 was used to identify a feasible set of sample points within the complete set of sample points. Sample points where the corresponding contact model outputs were outside their allowable bounds were deemed infeasible. First, initial surrogate contact models developed from a subset of the sample points were used to screen the complete set of sample points. Second, repeated static analyses were performed with the EF contact model to find physically realistic poses for the sample points passing the screening process. For each static analysis, the static contact loads in sensitive directions were compared to their corresponding applied values. A sample point was chosen only

when the maximum percent error was below 0.1%. Once all static analyses were completed and screened, contact loads corresponding to sensitive directions were fitted as a function of the six pose parameters, while contact loads corresponding to insensitive directions were fitted as a function of the contact loads in sensitive directions and pose parameters in insensitive ones.

Muscle and Contact Force Prediction

Once the surrogate contact models were incorporated into the knee model (Figure 4-3), muscle forces and associated TF and PF contact loads were predicted simultaneously using a two-level optimization approach. An outer level optimization (i.e., muscle force optimization) modified the muscle activations to minimize the cost function calculated by the inner level optimization (i.e., pose optimization) over all evaluated time frames. Given the initial guess for muscle activations, the pose optimization modified the nine free pose parameters (**q**) to to determine a static configuration that minimized the corresponding nine residual loads:

$$\min_{\mathbf{q}}\left(\left(F_{tibia\ SI}^{Residual}\right)^2+\left(F_{tibia\ ML}^{Residual}\right)^2+\left(T_{tibia\ VV}^{Residual}\right)^2+\sum_{i=1}^{3}\left(\left(F_{patella}^{Residual}\right)^2+\left(T_{patella}^{Residual}\right)^2\right)_i\right) \quad (4\text{-}1)$$

where $F_{tibia\ SI}^{Residual}$, $F_{tibia\ ML}^{Residual}$, and $T_{tibia\ VV}^{Residual}$ are the three tibial residual loads corresponding to superior-inferior translation, medial-lateral translation, and varus-valgus rotation, respectively, while $F_{patella}^{Residual}$ and $T_{patella}^{Residual}$ are patellar residual loads corresponding to three mutually perpendicular directions (i=1, 2, and 3). The tibial residual loads quantify the imbalance in inverse dynamics loads, muscle forces, and contact loads acting on the tibia, while the patellar residual loads quantify the imbalance in muscle forces and contact loads acting on the patella. The muscle force optimization modified the muscle activations to minimize one of two cost functions: 1) the sum of squares of muscle activations (Crowninshield and Brand, 1981; Forster et al., 1998; Anderson and Pandy, 2001) or 2) the sum of absolute values of the three

compressive loads (medial TF, lateral TF, and PF) (Schultz and Anderson, 1981), as shown below:

Minimize

$$\text{cost function 1}: \min_{\boldsymbol{a}} \sum_{i=1}^{11} a_i^2$$

or

$$\text{cost function 2}: \min_{\boldsymbol{a}} \left(\left| F_{\text{tibia SI}}^{\text{med contact}} \right| + \left| F_{\text{tibia SI}}^{\text{lat contact}} \right| + \left| F_{\text{patella SI}}^{\text{contact}} \right| \right)$$

Subject to

$$\text{constraints 1}: 0 \le a_i \le \infty$$

$$T_{\text{tibia FE}}^{\text{residual}} = 0 \qquad (4\text{-}2)$$

or

$$\text{constraints 2}: 0 \le a_i \le \infty$$

$$T_{\text{tibia FE}}^{\text{residual}} = 0$$

$$F_{\text{tibia AP}}^{\text{residual}} = 0$$

$$T_{\text{tibia IE}}^{\text{residual}} = 0$$

by varying muscle activation design variables $a_i \left(i = 1,...,11 \right)$

Each cost function was employed with two different constraint sets. Both constraint sets limited muscle activations to be positive and required the flexion-extension torque applied to the tibia by muscle and contact loads to match the flexion-extension torque from inverse dynamics. The first constraint set represents the approach commonly used in the literature. The second constraint set added matching of the anterior-posterior force and internal-external torque from inverse dynamics, thereby narrowing the feasible solution space further. Muscle activations were not constrained to be less than 1 (i.e., the physiological situation) to allow the optimizer to find feasible solutions that were not limited by upper bounds on muscle force production.

Eight optimization problems were formulated based on Equation 4-2 (Table 4-1). The problems consisted of all possible combinations of the two cost functions and two constraint sets, with each problem being solved with and without additional constraints (not shown in Equation

4-2) to match the in vivo medial and lateral contact force measurements. Problems that forced the solution to match the in vivo contact force measurements were termed "cheating" formulations (i.e., problems C1 through C4 in Table 4-1) while those that did not force the solution to match the in vivo measurements were termed "honest" formulations (i.e., problems H1 through H4 in Table 4-1). Problem C1 was used to evaluate the ability of a musculoskeletal model with surrogate contact to reproduce optimization predictions of muscle and contact forces generated using the same musculoskeletal model but with elastic foundation contact. Differences between the solutions generated by the two models were quantified by calculating maximum absolute errors (MAE) and root mean square errors (RMSE) in bone poses, contact loads, and muscle forces. The "cheating" formulations also allowed us to verify that the proposed musculoskeletal model was theoretically capable of matching the in vivo contact force measurements and to generate muscle force and PF contact force solutions that are likely to be within a physiological range. The "honest" formulations allowed us to evaluate how well different optimization problem formulations could predict the in vivo contact force measurements without knowing them. To keep computation time tractable, optimization solutions were generated at 5% increments across the gait cycle.

Results

The optimization performed with the surrogate contact models closely reproduced the 3D motions, contact forces, and muscle forces predicted with the EF contact models for problem formulation C1. On average, for both TF and PF joints, the RMSE for translations/rotations and forces/torques were less than 0.6 mm/ 0.8deg and 7 N/ 0.5 N-m, respectively (Table 4-2 and Table 4-3). The RMSE between surrogate-based and EF-based muscle force predictions were below 15 N (Table 4-4). While each optimization performed with the EF contact model required

approximately 90 minutes of CPU time, each optimization performed using the surrogate contact model required 2 minutes.

The medial and lateral tibial contact forces obtained from all "cheating" formulations but not all "honest" formulations were in good agreements with in vivo contact measurements (Figure 4-4). Overall, "honest" formulations predicted medial tibial contact forces that were comparable in magnitude and lateral tibial contact forces that were too low compared to measured tibial contact forces. Among other "honest" problem formulations, formulation H1 was able to obtain a better agreement between predicted and measured medial contact force throughout whole gait cycle while formulation H3 achieved the best agreement for the first 25% of the gait. For PF joint, the contact forces F_y obtained form eight problem formulations had the similar shape.

Muscle forces calculated from eight problem formulations were significantly different from each other (Figure 4-5). Compared to "honest" formulations, "cheating" formulations generated higher central and medial muscle forces and higher lateral muscle forces over the midstance and most of stance phase, respectively. None of the muscles exceed their maximum isometric forces except for BFSH and TFL.

Discussion

This study used a novel surrogate contact modeling approach to facilitate simultaneous prediction of muscle and contact forces in the knee during gait. The surrogate modeling approach involves replacing computationally costly contact models with computationally cheap contact models constructed using data points sampled from the original contact models. Once the surrogate models are constructed, they are used in place of the original models to eliminate a computational bottleneck caused by repeated contact analyses. Though the computational cost of constructing the three surrogate contact models is high (on the order of 30 hours of CPU time),

this time is quickly redeemed by the greatly reduced computational cost of each one cycle gait optimization (about 42 minutes of CPU time compared to 32 hours of CPU time for the EF contact models), especially when multiple gait optimizations are performed. The “cheating” problem formulations demonstrated that the proposed musculoskeletal model provided enough dimensionality in the magnitudes and directions of muscle forces to reproduce the in vivo contact force measurements. In addition, problem C1 demonstrated that the surrogate contact models could accurately reproduce the EF-based optimization results in a much shorter amount of CPU time. The "honest" formulations evaluated four different optimization problems and demonstrated that contact loads and muscle forces were significantly influenced by the optimization formulations, with different problem formulations matching medial contact force the best at different points in the gait cycle.

Simultaneous prediction of muscle and contact forces provides at least three advantages over sequential prediction, where muscle forces calculated without consideration of contact forces are subsequently used to calculate contact forces (Taylor et al., 2004; Kim et al., 2008). First, a simultaneous approach eliminates the need to make assumptions about the inverse dynamics loads to which contact forces do not contribute. By incorporating contact geometry into the model, one can satisfy all six inverse dynamics loads simultaneously at the tibiofemoral joint, and more loads than just the flexion-extension torque can be used as constraints. This benefit was demonstrated by problem formulations C2, C4, H2, and H4, where three inverse dynamics loads were used as equality constraints. The results from these four problems were different from the corresponding problems where only the flexion-extension torque served as a constraint, consistent with previous studies reporting that muscle force predictions are greatly

influenced by the number of inverse dynamics loads used as constraints (Buchanan et al., 1996; Glitsch et al., 1997; Li et al., 1998).

Second, a simultaneous approach makes it possible to utilize cost functions that are unrelated to muscles. This benefit was demonstrated by problems C3, C4, H3, and H4, where the cost function minimized compressive force in the three compartments of the knee rather than minimizing muscle activations, forces, stresses, or other muscle-related quantities. Minimization of compressive contact force predicted better medial contact force results in early stance phase than did minimization of muscle activations. This observation suggests that the body may use different cost functions (or different combinations of cost functions) at different points in the gait cycle to produce muscle and contact forces (Figure 4-4).

Third, a simultaneous approach allows one to match in vivo contact force measurements as additional constraints. Without these constraints, it would have been extremely difficult to evaluate whether our musculoskeletal model was theoretically capable of reproducing the in vivo medial and lateral contact force measurements. The addition of these constraints also made it possible to hypothesize whether the cheating solutions are representative of the in vivo situation. Whereas the predicted compressive forces on the patella were similar for all four cheating solutions (Figure 4-4), the predicted muscle forces were quite different (Figure 4-5). Thus, the predicted PF contact forces may provide reasonable estimates for the in vivo situation, while it is unclear which "cheating" formulation (if any) produced the most physiologically realistic muscle force predictions. This finding is not surprising given the amount of indeterminacy in the muscle force prediction problem. To gain confidence in the muscle force predictions, one would need to identify an "honest" problem formulation that consistently predicted the correct medial and lateral contact forces for a variety of gait motions. Muscle electromyographic data from the same

patient would also be valuable for evaluating the reasonableness of the predicted muscle activations. For the PF contact forces, the fact that the shapes (but not magnitudes) of the curves are similar to published studies that did not utilize contact models provides additional evaluation of these results (Ward and Power, 2004).

Our results highlight at least four possible musculoskeletal model inadequacies that may have contributed to the high errors observed in the predicted lateral contact forces. First, control of frontal plane stability may be critical for predicting lateral contact force accurately. "Honest" formulations may have generated low lateral contact force since the need for frontal plane stability was not accounted for in the problem formulation. It may be possible to increase frontal plane stability, and thus lateral contact force, by preferentially selecting muscles that contribute the most to varus-valgus stiffness (Nathan et al., 2008). Second, our musculoskeletal model only included a knee joint model, though some muscles were biarticular (e.g., TFL). Consequently, force-producing constraints imposed on biarticular muscles by neighboring joints were not account for in our problem formulations. In particular, for lateral biarticular muscles crossing the knee and hip (e.g., TFL), the need to balance the frontal plane moment at the hip to keep the pelvis level could contribute to increased lateral muscle force and hence lateral knee contact force. Third, the lack of lateral collateral ligament (LCL) in our musculoskeletal model may have affected the lateral contact force predictions, as the LCL is the primary structure resisting varus angulation of the knee (Gollehon et al., 1987; LaPrade et al., 2004). Clearly, adding an LCL with high stiffness to the model could increase lateral compressive load. To reproduce the in vivo lateral contact force measurements, the LCL would need to provide approximately 400 N of force throughout most of the gait cycle (Figure 4-13). However, this requirement is physiologically unrealistic as the mean failure strength for the LCL is approximately 450 N

(Ciccone et al., 2005). Furthermore, in vitro measurements indicate that lateral compartment contact force due to ligaments alone is on the order of only 50 to 100 N (Hawary et al., 2003). Fourth, no muscle physiology (i.e., force-length and force-velocity properties) and no activation or contraction dynamics were incorporated into our muscle model. Anderson et al. (2001b) reported that inclusion of those properties had only a small effect on muscle force predictions during gait, suggesting that the accuracy of the muscle forces predicted by optimization methods depends mainly on the accuracy of the inverse dynamics loads.

In summary, this study presented a way to predict muscle and contact forces simultaneously at the knee using surrogate contact modeling methods. The "cheating" optimization problem formulations revealed that our musculoskeletal model with surrogate contact was capable of reproducing in vivo medial and lateral contact force measurements and that the surrogate contact solutions were reliable. The "honest" formulations revealed that our musculoskeletal model could predict in vivo medial contact force measurements well but underestimated lateral contact force measurements. Demonstrating the feasibility of this simultaneous approach is a crucial step toward its application to a large number of trials or patients. Future studies will increase musculoskeletal model complexity by including control of frontal plane stability in the optimization formulation and by adding ligaments and other lower extremity joints.

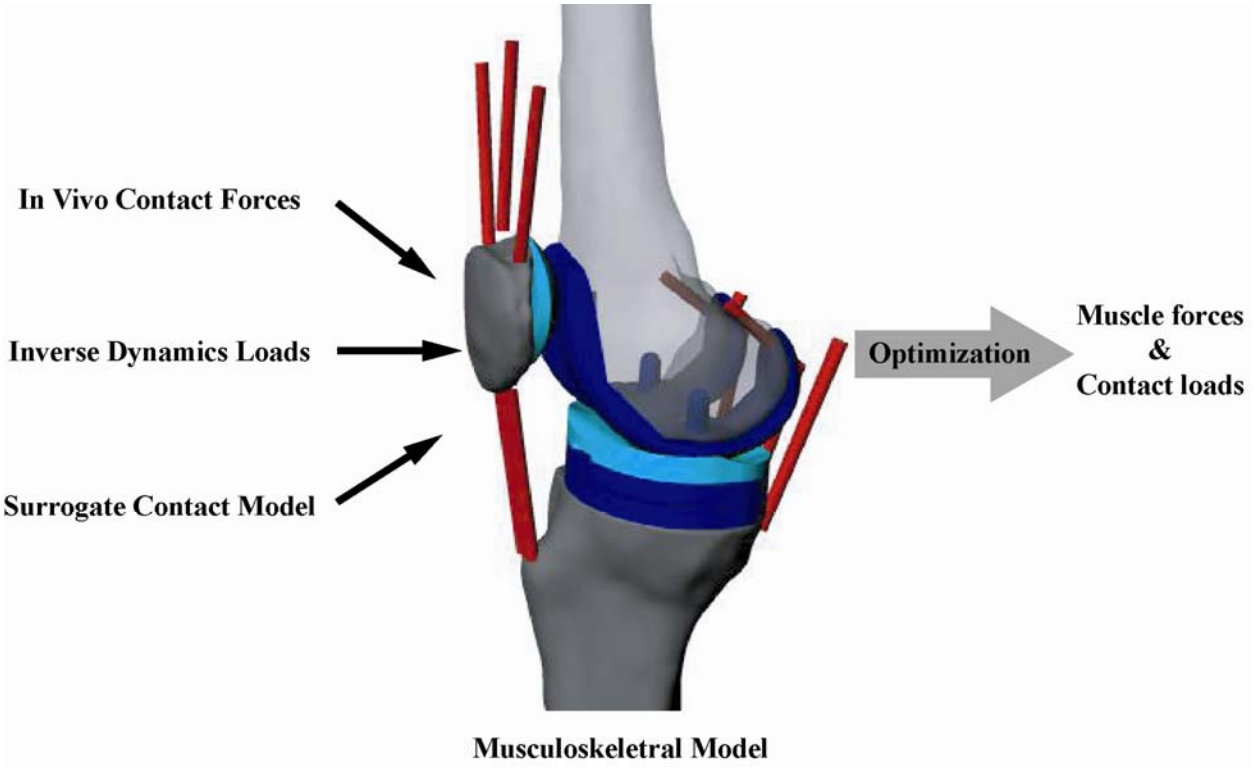

Figure 4-1. Overview of the process of using musculoskeletal model to predict muscle forces and contact loads simultaneously over the entire gait cycle. Surrogate contact model is created to provide computationally efficient contact analysis for musculoskeletal model during optimization process. The optimization is solved subject to constraints that the inverse dynamics loads were balanced out by the combination of predicted contact and muscle forces. The predictions are evaluated by comparing measured and predicted tibial contact forces.

Table 4-1. Eight different optimization problem formulations. C1 to C4 represent four "cheating" formulations while H1 to H4 represent four "honest" formulations.

	Constraints 1	Constraints 2
Cost Function 1	C1 (H1)	C2 (H2)
Cost Function 2	C3 (H3)	C4 (H4)

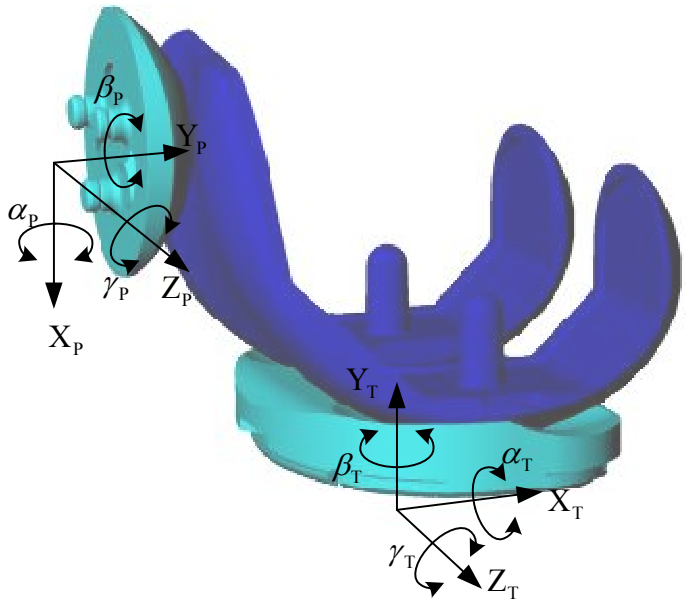

Figure 4-2. A 12 degree of freedom total knee replacement contact model. The femoral component is fixed to the ground while tibial insert and patella are connected to the femoral component via two 6 DOF joints (anterior-posterior translation X, superior-inferior translation Y, medial-lateral translation Z, varus-valgus rotation α, internal-external rotation β, and flexion-extension γ).

Table 4-2. Summary of root mean square and maximum absolute joint kinematics errors produced by surrogate contact model for problem formulation C1.

	TF joint		PF joint	
	RMSE	MAE	RMSE	MAE
X (mm)	0.001	0.002	0.300	0.882
Y (mm)	0.005	0.020	0.568	1.183
Z (mm)	0.017	0.042	0.452	0.785
α (deg)	0.007	0.024	0.792	2.113
β (deg)	0.001	0.003	0.133	0.356
γ (deg)	0.000	0.001	0.330	0.765

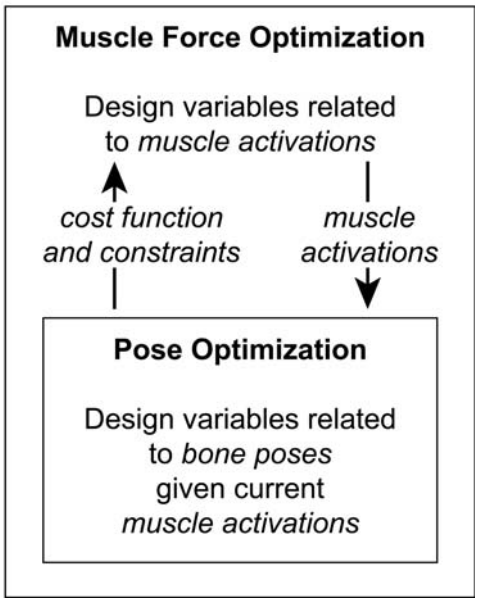

Figure 4-3. Simultaneous approache for predicting muscle forces and contact loads. The pose optimization performs repeated static analyses to equilibrate the free DOFs of the musculoskeletal model given the current muscle activations. The muscle force optimization adjusts the muscle activations to evaluate the cost function and constraints calculated by the pose optimization.

Table 4-3. Summary of root mean square and maximum absolute joint dynamics errors produced by surrogate contact model for problem formulation C1.

	TF joint				PF joint	
	Medial contact		Lateral contact			
	RMSE	MAE	RMSE	MAE	RMSE	MAE
Fx (N)	0.75	2.63	2.32	7.83	0.82	2.58
Fy (N)	0.49	1.88	1.91	8.54	6.42	16.50
Fz (N)	3.42	12.69	2.73	5.90	2.71	6.68
Tx (N-m)	0.14	0.40	0.35	1.30	0.05	0.15
Ty (N-m)	0.02	0.04	0.05	0.17	0.00	0.01
Tz (N-m)	0.03	0.06	0.08	0.25	0.04	0.11

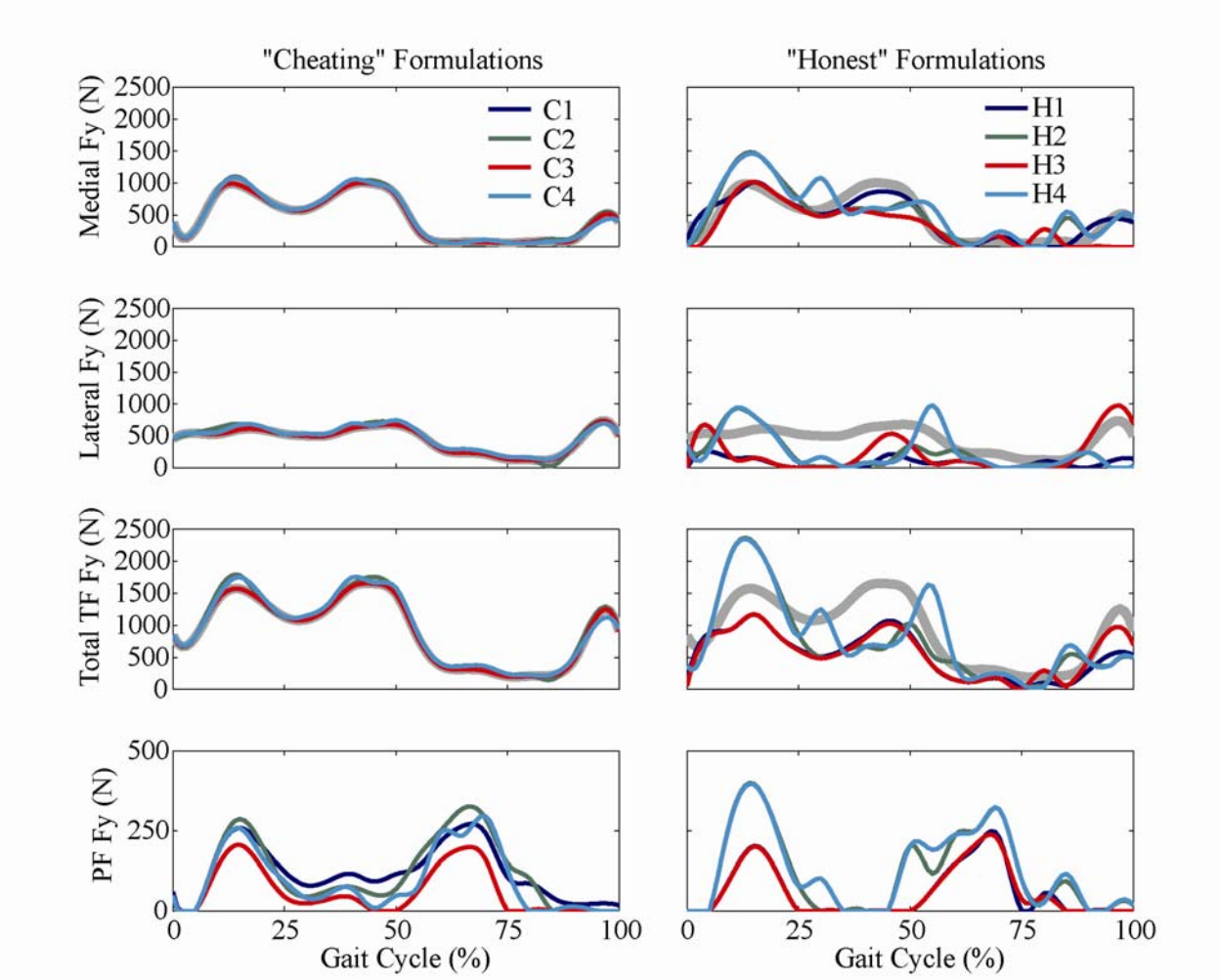

Figure 4-4. Comparison between the in vivo contact measurements (color in grey) and predictions from eight optimization problem formulations.

Table 4-4. Summary of root mean square and maximum absolute muscle force errors produced by surrogate contact model for problem formulation C1.

	Peak Isometric Force (N)	RMSE (N)	MAE (N)
RF	1319.84	11.53	46.92
VAS	6857.89	13.50	53.85
MHAMS	1941.46	7.07	19.45
MGAS	825.47	7.61	20.27
LGAS	825.47	10.94	25.16
BFLH	872.25	12.14	33.71
BFSH	228.26	7.97	26.90
TFL	262.16	7.54	18.56

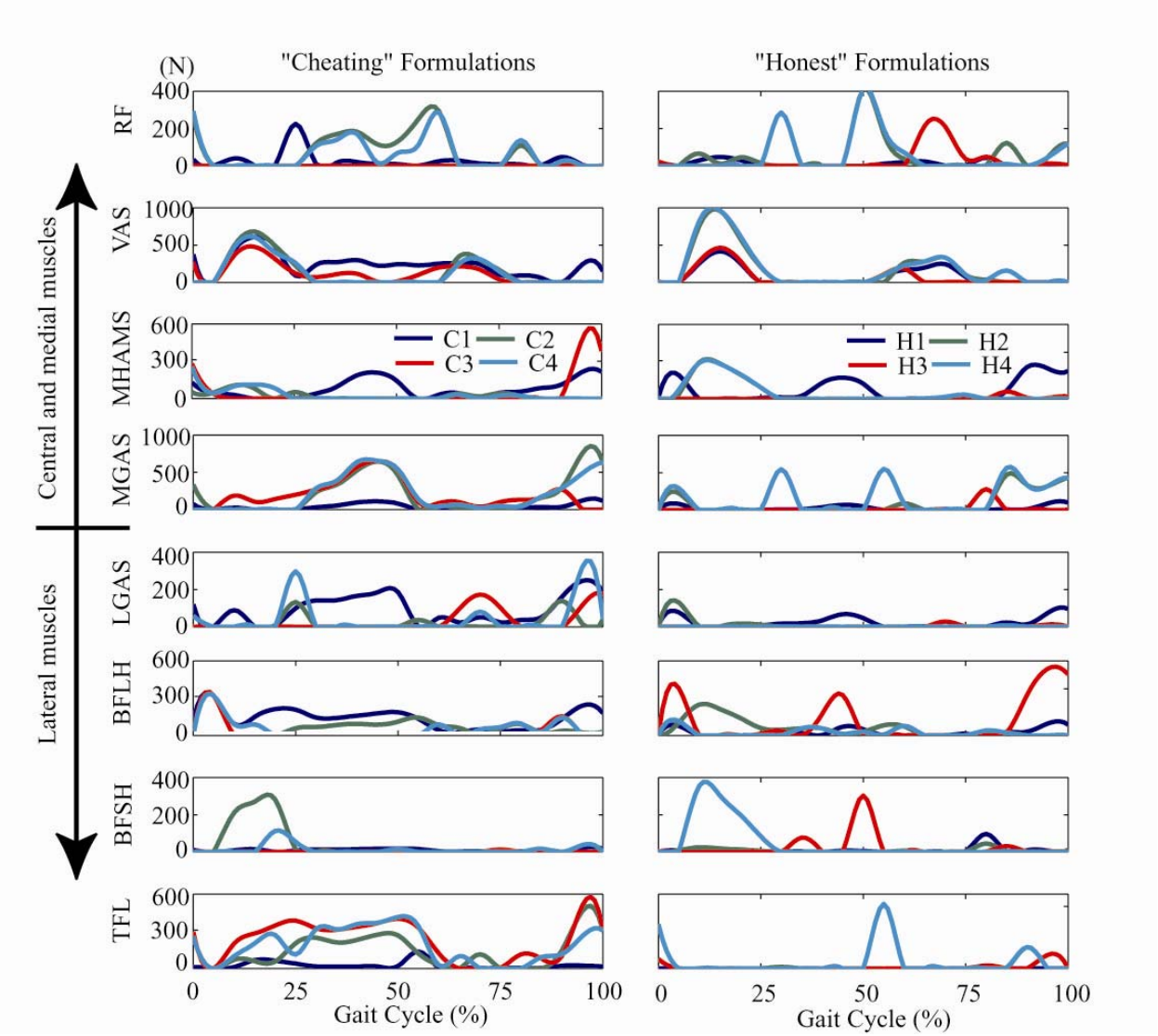

Figure 4-5. Predictions of muscle forces from eight optimization problem formulations.

CHAPTER 5
CONCLUSION

The main motivation of bringing surrogate modeling approach into biomechanics field is to efficiently predict the quantities that are difficult to measure in vivo (e.g., joint contact force, contact pressure, muscle force). Those quantities would be valuable for preventing and treating joint injuries and for improving the longevity of joint replacements. Although a variety of computational models have been used for this purpose, none of them is computationally efficient enough for studies involving repeated contact analysis.

Unlike other computational models, a surrogate modeling approach provides a way to eliminate contact-related computational bottlenecks for any type of simulations possessing one or more articular contacts. Our results first demonstrated that the surrogate modeling approach can accurately and efficiently reproduce either 2D or 3D dynamic simulations results originally generated by the elastic foundation contact model. The fact that a Stanmore knee simulator machine is highly sensitive to small variations in motion and load inputs is presented as a practical application of using surrogate contact model.

Given what we have learned from chapter 3, same technique could be applied to the similar analysis in chapter 2. Instead of using Hertzian contact theory and linear relationship between F_x and F_y, the concept of sensitive directions and reasonable design space approach could be used to generate a new accurate surrogate contact model for 2D dynamic contact simulation. The new surrogate should perform better than the original one since new approach will minimize the number of surrogates needed to calculate the contact forces.

Followed by the successful single joint simulation, we then demonstrated that same surrogate modeling approach can be applied to a multi-joint contact analysis for predicting muscle and contact forces simultaneously. The task was done by integrating a surrogate contact

model into a musculoskeletal model of the knee. The results showed that the proposed surrogate-based simultaneous approach could finish an optimization analysis within a practical CPU time to investigate the relationship between contact loads and muscle forces during gait. Although we have shown that the simultaneous approach could reproduce the in vivo contact measurements when they were used as additional constraints, further research is required to resolve the underestimation of lateral contact force.

Given the agreements of the surrogate-based results and EF-based results, a possible extension of current study could be to develop multiple surrogate contact models for all lower extremity joints other than just knee joint. Once finished, it would be possible to create a full-leg musculoskeletal model with multiple surrogate contact models that can be simulated in a short amount of CPU time. Another future topic involves the development of the surrogate modeling software. Once developed, the software along with a brief tutorial will be freely available throughout the internet. These resources will allow researchers to learn how the surrogate-based modeling approach can be used for their projects.

APPENDIX A
COMPUTATIONAL WEAR PREDICTION

The wear analysis calculated the depth δ of material removed from each element over N cycles based on Archard's wear law (Archard and Hirst, 1956)

$$\delta = Nk\sum_{i=1}^{n} p_i d_i = Nk\sum_{i=1}^{n} p_i \left|\mathbf{v}_i\right| \Delta t \tag{A1}$$

where k is the assumed material wear rate (1×10^{-7} mm^3/Nm; Fisher et al., 1994), i is a discrete time instant during the one-cycle dynamic simulation (1 through n), p_i is the contact pressure on the element at that instant, and d_i is the sliding distance experienced by the element, calculated as the product of magnitude $|\mathbf{v}_i|$ of the element's relative slip velocity and time increment Δt.

The wear volume was calculated by multiplying each element wear depth by the corresponding element A

$$V = \delta A = \left(Nk\sum_{i=1}^{n} p_i \left|\mathbf{v}_i\right| \Delta t \right) A \tag{A2}$$

Since $p_i A$ is just the normal contact force F_i applied to the element, the wear volume can be expressed concisely as

$$V = Nk\sum_{i=1}^{n} F_i \left|\mathbf{v}_i\right| \Delta t \tag{A3}$$

LIST OF REFERENCES

An, K.N., Himenso, S., Tsumura, H., Kawai, T., Chao, E.Y.S., 1990. Pressure distribution on articular surfaces: application to joint stability analysis. Journal of Biomechanics 23, 1013-1020.

Anderson, F.C., Pandy, M.G., 2001a. Dynamic optimization of human walking. Journal of Biomechanical Engineering-Transactions of the Asme 123, 381-390.

Anderson, F.C., Pandy, M.G., 2001b. Static and dynamic optimization solutions for gait are practically equivalent. Journal of Biomechanics 34, 153-161.

Archard, J.F., Hirst, W., 1956. The wear of metals under unlubricated conditions. Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences 236, 397-&.

Bai, B., Baez, J., Testa, N.N., Kummer, F.J., 2000. Effect of posterior cut angle on tibial component loading. Journal of Arthroplasty 15, 916-920.

Balabanov, V.O., Giunta, A.A., Golovidov, O., Grossman, B., Mason, W.H., Watson, L.T., Haftka, R.T., 1999. Reasonable design space approach to response surface approximation. Journal of Aircraft 36, 308-315.

Barnett, P.I., Fisher, J., Auger, D.D., Stone, M.H., Ingham, E., 2001. Comparison of wear in a total knee replacement under different kinematic conditions. Journal of Materials Science-Materials in Medicine 12, 1039-1042.

Bei, Y.H., Fregly, B.J., 2004. Multibody dynamic simulation of knee contact mechanics. Medical Engineering & Physics 26, 777-789.

Bitsakos, C., Kerner, J., Fisher, I., Amis, A.A., 2005. The effect of muscle loading on the simulation of bone remodeling in the proximal femur. Journal of Biomechanics 38, 133-139.

Blankevoort, L., Kuiper, J.H., Huiskes, R., Grootenboer, H.J., 1991. Articular contact in a three-dimensional model of the knee. Journal of Biomechanics 24, 1019-1031.

Blunn, G.W., Walker, P.S., Joshi, A., Hardinge, K., 1991. The dominance of cyclic sliding in producing wear in total knee replacements. Clinical orthopaedics and related research, 253-260.

Bottasso, C. L., Prilutsky, B. I., Croce, A., Imberti, E., Sartirana, S., 2006. A numerical procedure for inferring from experimental data the optimization cost functions using a multibody model of the neuromusculoskeletal system. Multibody System Dynamics 16, 123–154.

Bouzid, A.H., de Technologie Superieure, E., Champliaud, H., 1998. On the use of dual kriging interpolation for the evaluation of the gasket stress distribution in bolted joints. In Proceeding of the ASME Pressure Vessels and Piping Conference, 457, 77-83.

Box, G.E.P., Draper, N.R., 1987. Empirical model-building and response surfaces. Wiley, New York.

Brogliato, B., ten Dam, A.A., Paoli, L., Genot, F., Abadie, M., 2002. Numerical simulation of finite dimensional multibody nonsmooth mechanical systems. Applied Mechanics Reviews 55, 107.

Buchanan, T.S., Shreeve, D.A., 1996. An evaluation of optimization techniques for the prediction of muscle activation patterns during isometric tasks. Journal of Biomechanical Engineering 118, 565-574.

Buchanan, T.S., Lloyd, D.G., Manal, K., Besier, T.F., 2004. Neuromusculoskeletal modeling: estimation of muscle forces and joint moments and movements from measurements of neural command. Journal of Applied Biomechanics 20, 367-95.

Caruntu, D.I., Hefzy, M.S., 2004. 3-D anatomically based dynamic modeling of the human knee to include tibio-femoral and patello-femoral joints. Journal of Biomechanical Engineering 126, 44-53.

Challis, J.H., Kerwin, D.G., 1993. An analytical examination of muscle force estimations using optimization techniques. Proceedings of the Institution of Mechanical Engineers Part H 207, 139-148.

Chang, P.B., Williams, B.J., Santner, T.J., Notz, W.I., Bartel, D.L., 1999. Robust optimization of total joint replacements incorporating environmental variables. Journal of Biomechanical Engineering 121, 304-310.

Chao, E.Y.S., 2003. Graphic-based musculoskeletal model for biomechanical analyses and animation. Medical Engineering & Physics 25, 201-212.

Chen, S., Chng, E.S., Alkadhimi, K., 1996. Regularized orthogonal least squares algorithm for constructing radial basis function networks. International Journal of Control 64, 829-837.

Cheng, B., Titterington, D.M., 1994. Neural Networks - a Review from a Statistical Perspective. Statistical Science 9, 2-30.

Ciccone, W.J., Bratton, D.R., Weinstein, D.M., Walden, D.L., Elias, J.J., 2005. Structural properties of lateral collateral ligament reconstruction at the fibular head. The American Journal of Sports Medicine 34, 24-28.

Cohen, Z.A., Roglic, H., Grelsamer, R.P., Henry, J.H., Levine, W.N., Mow, V.C., Ateshian, G.A., 2001. Patellofemoral stresses during open and closed chain exercises – an analysis using computer simulation. The American Journal of Sports Medicine 29, 480-487.

Cox, S.E., Haftka, R.T., Baker, C.A., Grossman, B., Mason, W.H., Watson, L.T., 2001. A comparison of global optimization methods for the design of a high-speed civil transport. Journal of Global Optimization 21, 415-433.

Cressie, N., 1992. STATISTICS FOR SPATIAL DATA. Terra Nova 4, 613-617.

Crowninshield, R.D., Brand, R.A., 1981. A physiologically based criterion of muscle force prediction in locomotion. Journal of Biomechanics 14, 793-801.

Daffertshofer, A., Lamoth, C.J.C., Meijer, O.G., Beek, P.J., 2004. PCA in studying coordination and variability: a tutorial. Clinical Biomechanics 19, 415-428.

Delp, S.L., Arnold, A.S., Speers, R.S., Moore, C.A., 1996. Hamstrings and psoas lengths during normal and crouch gait: implications for muscle-tendon surgery. Journal of Orthopaedic Research 14, 144-151.

Dennerlein, J.T., Diao, E., Mote, C.D., Rempel, D.M., 1998. Tensions of the flexor digitorum superficialis are higher than a current model predicts. Journal of Biomechanics 31, 295-301.

DesJardins, J.D., Walker, P.S., Haider, H., Perry, J., 2000. The use of a force-controlled dynamic knee simulator to quantify the mechanical performance of total knee replacement designs during functional activity. Journal of Biomechanics 33, 1231-1242.

Diwekar, U.M., 2003. Introduction to Applied Optimization. Springer.

D'Lima, D.D., Hermida, J.C., Chen, P.C., Colwell, C.W., 2001. Polyethylene wear and variations in knee kinematics. Clinical Orthopaedics and Related Research, 124-130.

D'Lima, D.D., Townsend, C.P., Arms, S.W., Morris, B.A., Colwell, C.W., 2005. An implantable telemetry device to measure intra-articular tibial forces. Journal of Biomechanics 38, 299-304.

D'Lima, D.D., Patil, S., Steklov, N., Slamin, J.E., Colwell, C.W., 2006. Tibial forces measured in vivo after total knee arthroplasty. Journal of Arthroplasty 21, 255-262.

D'Lima, D.D., Patil, S., Steklov, N., Chien, S., Colwell, C.W., 2007. In vivo knee moments and shear after total knee arthroplasty. Journal of Biomechanics 40, S11-S17.

Elias, J.J., Wilson, D.R., Adamson, R., Cosgarea, A.J., 2004. Evaluation of a computational model used to predict the patellofemoral contact pressure distribution. Journal of Biomechanics 37, 295-302.

Erdemir, A., McLean, S., Herzog, W., van den Bogert, A.J., 2007. Model-based estimation of muscle forces exerted during movements. Clin Biomech (Bristol, Avon) 22, 131-54.

Fang, H., Rais-Rohani, M., Liu, Z., Horstemeyer, M.F., 2005. A comparative study of metamodeling methods for multiobjective crashworthiness optimization. Computers & Structures 83, 2121-2136.

Finni, T., Komi, P.V., Lukkariniemi, J., 1998. Achilles tendon loading during walking: application of a novel optic fiber technique. European Journal of Applied Physiology 77, 289-91.

Fishman, G.S., 1996. Monte Carlo: Concepts, Algorithms, and Applications. Springer.

Forster, E., Simon, U., Augat, P., and Claes, L., 2004. Extension of a state-of-the-art optimization criterion to predict co-contraction. Journal of Biomechanics 37, 577-581.

Fregly, B.J., Bei, Y.H., Sylvester, M.E., 2003. Experimental evaluation of an elastic foundation model to predict contact pressures in knee replacements. Journal of Biomechanics 36, 1659-1668.

Fregly, B.J., Sawyer, W.G., Harman, M.K., Banks, S.A., 2005. Computational wear prediction of a total knee replacement from in vivo kinematics. Journal of Biomechanics 38, 305-314.

Fregly, B.J., Reinbolt, J.A., Rooney, K.L., Mitchell, K.H., Chmielewski, T.L., 2007. Design of patient-specific gait modifications for knee osteoarthritis rehabilitation. IEEE Transactions on Biomedical Engineering 54, 1687-1695.

Fregly, B.J., Banks, S.A., D'Lima, D.D., Colwell, C.W., Jr., 2008. Sensitivity of knee replacement contact calculations to kinematic measurement errors. Journal of Orthopaedic Resarch 26, 1173-9.

Géradin, M., Cardona, A., 2001. Flexible multibody dynamics : a finite element approach. John Wiley, Chichester ; New York.

Gerstmayr, J., Schoberl, J., 2006. A 3D finite element method for flexible multibody systems. Multibody System Dynamics 15, 309-324.

Gibson, N., Graham, H.K., Love, S., 2007. Botulinum toxin A in the management of focal muscle overactivity in children with cerebral palsy. Disability and Rehabilitation 29, 1813-1822.

Giddings, V.L., Kurtz, S.M., Edidin, A.A., 2001. Total knee replacement polyethylene stresses during loading in a knee simulator. Journal of Tribology 123, 842-847.

Girosi, F., 1998. An equivalence between sparse approximation and support vector machines. Neural Computation 10, 1455-1480.

Giunta, A.A., Balabanov, V., Burgee, S., Grossman, B. Haftka, R.T., Mason, W.H., Watson, L.T., 1997. Multidisciplinary optimisation of a supersonic transport using design of experiments theory and response surface modeling. Aeronautical Journal 101, 347-356.

Glitsch, U., Baumann, W., 1997. The three-dimensional determination of internal loads in the lower extremity. Journal of Biomechanics 30, 1123-1131.

Glocker, C., 1999. Formulation of spatial contact situations in rigid multibody systems. Computer Methods in Applied Mechanics and Engineering 177, 199-214.

Godest, A.C., Meaugonin, M., Haug, E., Taylor, M., Gregson, P.J., 2002. Simulation of a knee joint replacement during a gait cycle using explicit finite element analysis. Journal of Biomechanics 35, 267-276.

Gollehon, D.L., Tozilli, P.A., Warren, R.F., 1987. The role of the posterolateral and cruciate ligaments in the stability of the human knee: a biomechanical study. The Journal of Bone and Joint Surgery 69, 233-242.

Gupta, A., Ding, Y., Xu, L., Reinikainen, T., 2006. Optimal parameter selection for electronic packaging using sequential computer simulations. Journal of Manufacturing Science and Engineering 128, 705-715.

Halloran, J.P., Easley, S.K., Petrella, A.J., Rullkoetter, P.J., 2005a. Comparison of deformable and elastic foundation finite element simulations for predicting knee replacement mechanics. Journal of Biomechanical Engineering 127, 813-818.

Halloran, J.P., Petrella, A.J., Rullkoetter, P.J., 2005b. Explicit finite element modeling of total knee replacement mechanics. Journal of Biomechanics 38, 323-331.

Hammersley, J.M., 1960. Related Problems .3. Monte-Carlo Methods for Solving Multivariable Problems. Annals of the New York Academy of Sciences 86, 844-874.

Hawary R.E., Roth, S.E., Harwood, J.C., Johnson, J.A., King, G.J.W., Chess, D.G., 2003. Ligamentous balancing in total knee arthroplasty: an in-vitro load cell analysis. Journal of Bone and Joint Surgery - British Volume 90.

Higginson, J.S., Zajac, F.E., Neptune, R.R., Kautz, S.A., Delp, S.L., 2006. Muscle contributions to support during gait in an individual with post-stroke hemiparesis. J Biomech 39, 1769-77.

Hippmann, G., 2004. An algorithm for compliant contact between complexly shaped bodies. Multibody System Dynamics 12, 345-362.

Hosder, S., Watson, L.T., Grossman, B., Mason, W.H., Kim, H., Haftka, R.T., Cox, S.E., 2001. Polynomial Response Surface Approximations for the Multidisciplinary Design Optimization of a High Speed Civil Transport. Optimization and Engineering 22, 431-452.

Hu, N., 1997. A solution method for dynamic contact problems. Computers & Structures 63, 1053-1063.

Hurwitz, D.E., Ryals, A.R., Block, J.A., Sharma, L., Schnitzer, T.J., Andriacchi, T.P., 2000. Knee pain and joint loading in subjects with osteoarthritis of the knee. Journal of Orthopaedic Research 18, 572-579.

Ishikawa, M., Komi, P.V., Grey, M.J., Lepola, V., Bruggemann, G.P., 2005. Muscle-tendon interaction and elastic energy usage in human walking. Journal of Applied Physiology 99, 603-608.

Jin, R., Chen, W., Simpson, T.W., 2001. Comparative studies of metamodelling techniques under multiple modelling criteria. Structural and Multidisciplinary Optimization 23, 1-13.

Jinha, A., Ait-Haddou, R., Herzog, W., 2006. Predictions of co-contraction depend critically on degrees-of-freedom in the musculoskeletal model. Journal of Biomechanics 39, 1145-1152.

Johnson, T.S., Laurent, M.P., Yao, J.Q., Gilbertson, L.N., 2001. The effect of displacement control input parameters on tibiofemoral prosthetic knee wear. Wear 250, 222-226.

Jones, D.R., Schonlau, M., Welch, W.J., 1998. Efficient global optimization of expensive black-box functions. Journal of Global Optimization 13, 455-492.

Kalagnanam, J.R., Diwekar, U.M., 1997. An efficient sampling technique for off-line quality control. Technometrics 39, 308-319.

Kane, T.R., Levinson, D.A., 1985. Dynamics, theory and applications. McGraw-Hill, New York.

Kaufman, K.R., An, K.N., Litchy, W.J., Chao, E.Y.S., 1991. Physiological Prediction of Muscle Forces .2. Application to Isokinetic Exercise. Neuroscience 40, 793-804.

Kaufman, M.D., Balabanov, V., Burgee, S.L., Giunta, A.A., Grossman, B., Haftka, R.T., Mason, W.H., and Watson, L.T., 1996. Variable-complexity response surface approximations for wing structural weight in HSCT design. Computational Mechanics 18, 112-126.

Kawanabe, K., Clarke, I.C., Tamura, J., Akagi, M., Good, V.D., Williams, P.A., Yamamoto, K., 2001. Effects of A-P translation and rotation on the wear of UHMWPE in a total knee joint simulator. Journal of Biomedical Materials Research 54, 400-406.

Khuri, A.I., Cornell, J.A., 1996. Response surfaces : designs and analyses. Marcel Dekker, New York.

Kim, H.J, Fernandez, J.W., Akbarshahi, M., Walter, J.P., Fregly, B.J., Pandy. M.G., 2008. Evaluation of predicted knee-joint muscle forces during gait using an instrumented knee implant. Journal of Orthopaedic Research (in press).

Knight, L.A., Pal, S., Coleman, J.C., Bronson, F., Haider, H., Levine, D.L., Taylor, M., Rullkoetter, P.J., 2007. Comparison of long-term numerical and experimental total knee replacement wear during simulated gait loading. Journal of Biomechanics 40, 1550-1558.

Koch, P.N., Simpson, T.W., Allen, J.K., Mistree, F., 1999. Statistical approximations for multidisciplinary design optimization: The problem of size. Journal of Aircraft 36, 275-286.

Kurtaran, H., Eskandarian, A., Marzougui, D., Bedewi, N.E., 2001. Crashworthiness design optimization using successive response surface approximations. Computational Mechanics 29, 409-421.

Kwak, S.D., Blankevoort, L., Ateshian, G.A., 2000. A mathematical formulation for 3D quasi-static multibody models of diarthrodial joints. Computer Methods in Biomechanics and Biomedical Engineering 3, 41-64.

LaPrade, R.F., Tso, A., Wentorf, F.A., 2004. Force measurements on the fibular collateral ligament, popliteofibular ligament, and popliteus tendon to applied loads. American Journal of Sports Medicine 32, 1695-1701.

Li, B., Melkote, S.N., 1999. An elastic contact model for the prediction of workpiece-fixture contact forces in clamping. Journal of Manufacturing Science and Engineering 121, 485-493.

Li, G., Sakamoto, M., Chao, E.Y.S., 1997. A comparison of different methods in predicting static pressure distribution in articulating joints. Journal of Biomechanics 30, 635-638.

Li, G., Kaufman, K.R., Chao, E.Y.S., Rubash, H.E., 1999. Prediction of antagonistic muscle forces using inverse dynamic optimization during flexion extension of the knee. Journal of Biomechanical Engineering 121, 316-322.

Lin, Y.C., Farr, J., Carter, K., Fregly, B.J., 2006. Response surface optimization for joint contact model evaluation. Journal of Applied Biomechanics 22, 120-130.

Liu, B., Haftka, R.T., Akgun, M.A., 2000. Two-level composite wing structural optimization using response surfaces. Structural and Multidisciplinary Optimization 20, 87-96.

Lophaven, S.N., Nielsen, H.B., Søndergaard, J., 2002. DACE-A Matlab Kriging Toolbox. Technical Report IMM-TR-2002-13, Informatics and Mathematical Modelling, Technical University of Denmark.

Mack, Y., Goel, T., Shyy, W., Haftka, R.T., 2007. Surrogate Model-Based Optimization Framework: A Case Study in Aerospace Design. Studies in Computational Intelligence 51, 323-342.

Mancuso, C.A., Sculco, T.P., Wickiewicz, T.L., Jones, E.C., Robbins, L., Warren, R.F., Williams-Russo, P., 2001. Patients' expectations of knee surgery. Journal of Bone and Joint Surgery-American Volume 83A, 1005-1012.

Martens, H., Naes, T., 1984. Multivariate Calibration. Chemometrics, Mathematics and Statistics in Chemistry.

Milner, C.E., Hamill, J., Davis, I., 2007. Are knee mechanics during early stance related to tibial stress fracture in runners? Clinical Biomechanics 22, 697-703.

Moran, M.F., Bhimji, S., Racanelli, J., Piazza, S.J., 2008. Computational assessment of constraint in total knee replacement. Journal of Biomechanics 41, 2013-20.

Myers, R.H., Montgomery, D.C., 1995. Response surface methodology : process and product in optimization using designed experiments. Wiley, New York.

Bunderson, N.E., Burkholder, T.J., Ting, L.H., 2008. Reduction of neuromuscular redundancy for postural force generation using an intrinsic stability criterion. Journal of Biomechanics 41,1537-1544.

Neptune, R.R., Wright, I.C., van den Bogert, A.J., 2000. Influence of orthotic devices and vastus medialis strength and timing on patellofemoral loads during running. Clinical Biomechanics 15, 611-618.

Noble, P.C., Gordon, M.J., Weiss, J.M., Reddix, R.N., Conditt, M.A., Mathis, K.B., 2005. Does total knee replacement restore normal knee function? Clinical Orthopaedics and Related Research, 157-165.

Pal, S., Haider, H., Laz, P.J., Knight, L.A., Rullkoetter, P.J., 2008. Probabilistic computational modeling of total knee replacement wear. Wear 264, 701-707.

Pandy, M.G., Sasaki, K., and Kim, S., 1997. A three-dimensional musculoskeletal model of the human knee joint. Part 1: Theoretical construction. Computer Methods in Biomechanics and Biomedical Engineering 1, 87-108.

Pandy, M.G., Sasaki, K., Kim, S., 1998. A three-dimensional musculoskeletal model of the human knee joint. part 1: theoretical construct. Computer Methods in Biomechanics and Biomedical Engineering 1, 87-108.

Piazza, S.J., Delp, S.L., 2001. Three-dimensional dynamic simulation of total knee replacement motion during a step-up task. Journal of Biomechanical Engineering 123, 599-606.

Pierce, J.E., Li, G., 2005. Muscle forces predicted using optimization methods are coordinate system dependent. Journal of Biomechanics 38, 695-702.

Potra, F.A., Anitescu, M., Gavreal, B., Trinkle, J., 2006. A linearly implicit trapezoidal method for integrating stiff multibody dynamics with contact, joints, and friction. International Journal for Numerical Methods in Engineering 66, 1079-1124.

Queipo, N.V., Pintos, S., Rincon, N., Contreras, N., Colmenares, J., 2002. Surrogate modeling-based optimization for the integration of static and dynamic data into a reservoir description. Journal of Petroleum Science and Engineering 35, 167-181.

Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Tucker, P.K., 2005. Surrogate-based analysis and optimization. Progress in Aerospace Sciences 41, 1-28.

Rawlinson, J.J., Furman, B.D., Li, S., Wright, T.M., Bartel, D.L., 2006. Retrieval, experimental, and computational assessment of the performance of total knee replacements. Journal of Orthopaedic Research 24, 1384-1394.

Reeves, N.D., Maganaris, C.N., Narici, M.V., 2003. Effect of strength training on human patella tendon mechanical properties of older individuals. Journal of Physiology-London 548, 971-981.

Reinbolt, J.A., Haftka, R.T., Chmielewski, T.L., Fregly, B.J., 2008. A computational framework to predict post-treatment outcome for gait-related disorders. Medical Engineering & Physics 30, 434-443.

Reinbolt, J.A., Schutte, J.F., Fregly, B.J., Koh, B.I., Haftka, R.T., George, A.D., Mitchell, K. H., 2005. Determination of patient-specific multi-joint kinematic models through two–level optimization. Journal of Biomechanics 38, 621–662.

Roux, W.J., Stander, N., Haftka, R.T., 1998. Response surface approximations for structural optimization. International Journal for Numerical Methods in Engineering 42, 517-534.

Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P., 1989. Design and analysis of computer experiments. Statistical Science 4, 409-435.

Santamaria, J., Vadillo, E.G., Gomez, J., 2006. A comprehensive method for the elastic calculation of the two-point wheel-rail contact. Vehicle System Dynamics 44, 240-250.

Schmidt, D.J., Arnold, A.S., Carroll, N.C., Delp, S.L., 1999. Length changes of the hamstrings and adductors resulting from derotational osteotomies of the femur. Journal of Orthopaedic Research 17, 279-285.

Schuind, F., Garcia-Elias, M., Cooney, W.P., 3rd, An, K.N., 1992. Flexor tendon forces: in vivo measurements. Journal of Hand Surgery 17, 291-298.

Schultz, A.B., Anderson, G.B.J., 1981. Analysis of loads on the lumbar spine. Spine 6, 76–82.

Sharf, I., Zhang, Y.N., 2006. A contact force solution for non-colliding contact dynamics simulation. Multibody System Dynamics 16, 263-290.

Simpson, T.W., Peplinski, J.D., Koch, P.N., Allen, J.K., 2001. Metamodels for computer-based engineering design: survey and recommendations. Engineering with Computers 17, 129-150.

Smith, M., 1993. Neural Networks for Statistical Modeling. John Wiley & Sons, Inc. New York, NY, USA.

Stansfield, B.W., Nicol, A.C., Paul, J.P., Kelly, I.G., Graichen, F., Bergmann, G., 2003, Direct comparison of calculated hip joint contact forces with those measured using instrumented implants. An evaluation of a three-dimensional mathematical model of the lower limb. Journal of biomechanics 36, 929-936.

Stewart, D.E., 2000. Rigid-body dynamics with friction and impact. Siam Review 42, 3-39.

Streator, J.L., 2003. Dynamic contact of a rigid sphere with an elastic half-space: A numerical simulation. Journal of Tribology 125, 25-32.

Sujith, O.K., 2008. Functional electrical stimulation in neurological disorders. European Journal of Neurology 15, 437-444.

Taylor, W., Heller, M., Bergmann, G., Duda, G., 2004. Tibio-femoral loading during human gait and stair climbing. Journal of Orthopaedic Research 22, 625-632.

Valero-Cuevas, F.J., Johanson, M.E., Towles, J.D., 2003. Towards a realistic biomechanical model of the thumb: the choice of kinematic description may be more critical than the solution method or the variability/uncertainty of musculoskeletal parameters. Journal of Biomechanics 36, 1019-1030.

van den Bogert, A.J., Smith, G.D., Nigg, B.M., 1994. *In vivo* determination of the anatomical axes of the ankle joint complex: an optimization approach. Journal of Biomechanics 27:1477–1488.

Vapnik, V.N., 1998. Statistical learning theory. John Wiley & Sons.

Wahba, G., 1987. Spline models for observational data. CBMS-NSF Regional Conference Series in Applied Mathematics, Based on a series of 10 lectures at Ohio State University at Columbus, March 23-27, 1987, Philadelphia: Society for Industrial and Applied Mathematics, 1990.

Walker, P.S., Blunn, G.W., Broome, D.R., Perry, J., Watkins, A., Sathasivam, S., Dewar, M.E., Paul, J.P., 1997. A knee simulating machine for performance evaluation of total knee replacements. Journal of Biomechanics 30, 83-89.

Wang, G.G., Shan, S., 2007. Review of metamodeling techniques in support of engineering design optimization. Journal of Mechanical Design 129, 370-380.

Ward, S.R., Powers, C.M., 2004. The influence of patella alta on patellofemoral joint stress during normal and fast walking. Clinical Biomechanics 19, 1040-1047.

Weiss, J.M., Noble, P.C., Conditt, M.A., Kohl, H.W., Roberts, S., Cook, K.F., Gordon, M.J., Mathis, K.B., 2002. What functional activities are important to patients with knee replacements? Clinical Orthopaedics and Related Research, 172-188.

Wendland, H., 1995. Piecewise polynomial, positive definite and compactly supported radial basis functions of minimal degree. Advances in Computational Mathematics 4, 389-396.

Wilson, D.R., Feikes, J.D., O'Connor, J.J., 1998. Ligaments and articular contact guide passive knee flexion. Journal of Biomechanics 31, 1127-36.

Wu, Y.T., Millwater, H.R., Cruse, T.A., 1990. Advanced probabilistic structural-analysis method for implicit performance functions. AIAA Journal 28, 1663-1669.

Zhao, D., Banks, S.A., D'Lima, D.D., Colwell, C.W., Fregly, B.J., 2007. In vivo medial and lateral tibial loads during dynamic and high flexion activities. Journal of Orthopaedic Research 25, 593-602.

Zhao, D., Banks, S.A., Mitchell, K.H., D'Lima, D.D., Colwell, C.W., Fregly, B.J., 2007. Correlation between the knee adduction torque and medial contact force for a variety of gait patterns. Journal of Orthopaedic Research 25, 789-797.

Zhao, D., Sakoda, H., Sawyer, W.G., Banks, S.A., Fregly, B.J., 2008. Predicting knee replacement damage in a simulator machine using a computational model with a consistent wear factor. Journal of Biomechanical Engineering 130, 011004.

BIOGRAPHICAL SKETCH

Yi-Chung Lin was born in Tainan, a city in southern Taiwan, in 1976. During his childhood, he enjoyed assembling plastic models. He has assembled many different varieties of plastic models (aircraft, automobile, robot), and gradually found himself a big fan of robots. This hobby eventually led him into the Department of Mechanical Engineering of the National Cheng Kung University, Taiwan. He received his Bachelor of Science in 1998 and joined the Army to fulfill the military service obligation. After receiving a Honorable Discharge, he went overseas and enrolled in the master's program in the Department of Mechanical and Aerospace Engineering at the University of Florida. During his master's program, he extended his interest from robots to the human body and joinined the Computational and Biomechanics Laboratory of Associate Professor B. J. Fregly in fall 2002. In 2004, Yi-Chung received a Master of Science degree in mechanical engineering with focus on biomechanics. At that time, he decided to stay in the same lab to pursue the Doctor of Philosophy degree. Yi-Chung would like to continue his research in biomechanics and to improve the quality of human daily life in the future.