Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
2021, Vol. 30(6) 1523–1537


Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09622802211003620
journals.sagepub.com/home/smm
SAGE

Calibration of surgical tools using multilevel modeling with LINEX loss function: Theory and experiment

Parisa Azimaee[1,2], Mohammad Jafari Jozani[2] and Yaser Maddahi[3]

Abstract

Quantifying the tool–tissue interaction forces in surgery can be used in the training process of novice surgeons to help them better handle surgical tools and avoid exerting excessive forces. A significant challenge concerns the development of proper statistical learning techniques to model the relationship between the true force exerted on the tissue and several outputs read from sensors mounted on the surgical tools. We propose a nonparametric bootstrap technique and a Bayesian multilevel modeling methodology to estimate the true forces. We use the linear exponential loss function to asymmetrically penalize the over and underestimation of the applied forces to the tissue. We incorporate the direction of the force as a group factor in our analysis. A weighted approach is used to account for the nonhomogeneity of read voltages from the surgical tool. Our proposed Bayesian multilevel models provide estimates that are more accurate than those under the maximum likelihood and restricted maximum likelihood approaches. Moreover, confidence bounds are much narrower and the biases and root mean squared errors are significantly smaller in our multilevel models with the linear exponential loss function.

Keywords

Calibration, bootstrap technique, multilevel modeling, Bayesian approach, microsurgery, medical tool, tool–tissue interaction force, surgical training

1 Introduction

Quantification of the force exerted during the performance of surgical operations is of importance in the training process of novice surgeons and surgical residents. A significant amount of errors in neurosurgical operations (more than 50%) is due to excessive force to the brain tissue.[1] Quantification of such forces would help trainees learn about safe margins of forces when dealing with human tissue, and help residents acquire surgical skills in practice. Knowledge of the force values may also significantly improve the learning curve, since the residents can "learn-by-doing", instead of only "observing" experienced surgeons performing surgical tasks.[2]

The development of educational systems, that precisely reflect the mechanical components, such as tool–tissue interaction forces, during an operation, is of importance. Without such systems, trainees require years of experience and multiple pre-clinical and clinical trials to become educated in dealing appropriately with human tissues. Therefore, there is an increasing demand to enhance the efficiency of the learning process and provide trainees with quantitative tools to assess their surgical skills.[3,4] This paper concerns microsurgical operations, where the

[1]Wawanesa Insurance, Winnipeg, MB, Canada
[2]Department of Statistics, University of Manitoba, Winnipeg, MB, Canada
[3]Tactile Robotics, Winnipeg, MB, Canada

Corresponding author:
Mohammad Jafari Jozani, Department of Statistics, University of Manitoba, Winnipeg, MB R3T 3Z2, Canada.
Email: m_jafari_jozani@umanitoba.ca

goal is to increase the accuracy and resolution of the functionality of the sensorized surgical forceps by adding some miniature force-sensing components called strain gauge.[2] The force exerted to the brain tissue generates a change in the electrical resistance of the strain gauge, and as a result, the voltage would change.[1–3,5] There is a relationship between the force applied to the brain tissue and the voltages read from the mounted strain gauges on the prongs of the surgical forceps. It is required to appropriately model the output voltages (*response variable*) and the force components (*explanatory variable*), to estimate the resultant force.

In this paper, we study a dataset that was obtained from a calibration station developed to calibrate a bipolar forceps instrumented with small strain gauges. The data set was provided by the Project neuroArm at the University of Calgary, and was partially used in our recent work in which the nonparametric bootstrap technique was employed to estimate two-dimensional (2D) forces exerted on a cadaveric brain during the performance of some neurosurgical tasks.[6] We developed a *Naïve* method using the underlying deterministic and physical properties pertinent to the strain gauges.[2] However, the *Naïve* approach has important limitations, as it does not allow to construct necessary precision measures and confidence intervals for the predicted forces. To overcome these problems, we developed another approach based on a nonparametric bootstrap technique to accurately estimate the amount of force associated with the observed signals in a 2D environment where the forces are implemented in only x and y directions. Necessary confidence intervals were developed and later used in the calibration of the sensorized bipolar forceps.[6]

In practice, forces are applied to the tissue in three directions. One needs to provide accurate estimates of forces along x-, y-, and z-axes using voltage signals read from three strain gauges. The first two strain gauges (measuring the deflections in x and y directions) were mounted onto the surface of each prong of the forceps. To place the third sensor for measuring the deflection along z-axis, the tool was altered and the strain gauge was installed inside a U-shaped notch created during alteration (see the inset in Figure 1).

Due to the nonlinear behavior of the strain gauges in the revised form of sensorized bipolar forceps, the signals measured from some of the strain gauges are often very noisy and out of range. As a result, the current approaches can provide accurate estimates of the forces in only x and y directions, and are not able to obtain reasonable results when data on the x, y, and z directions are used simultaneously. In this paper, we develop different techniques to address this important issue. To this end, we first adopt the nonparametric bootstrap technique for our 3D dataset and demonstrate that this technique does not properly estimate the interaction force in all three

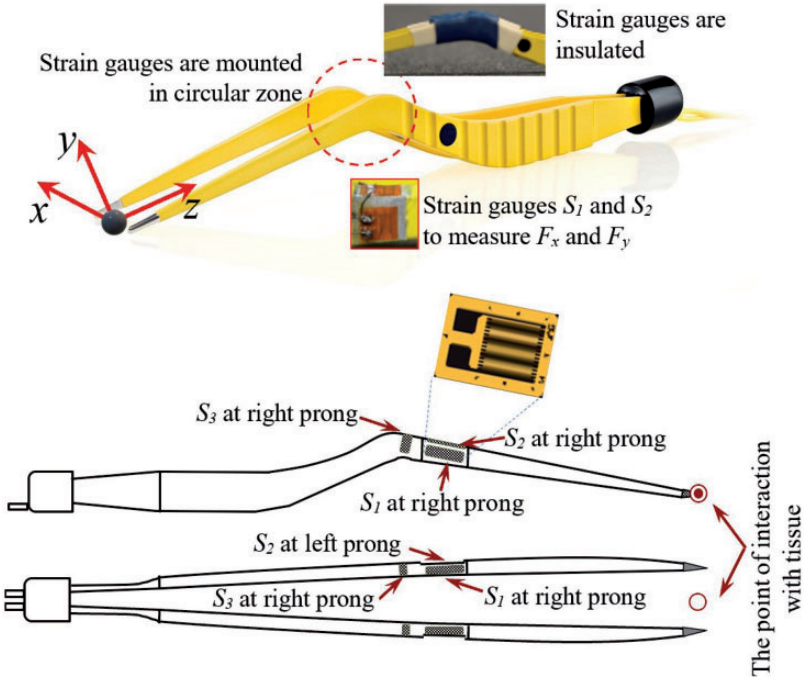

Figure 1. Schematic of the developed bipolar forceps that uses three strain gauges S_1, S_2, and S_3 to quantify the force components in x, y, and z directions, respectively. The bipolar forceps is covered with a nonconductive material to protect the circuit during the operation and sterilization.

directions. We then propose different approaches based on a weighted least squares regression[7] method as well as Multilevel modeling[8] using a Bayesian approach.[9] As underestimating the implemented forces is not desirable in our application, we develop a new Bayesian multilevel modeling approach using an asymmetric penalty (loss) function that penalizes underestimation more than overestimation. Theoretical results are augmented with a real data analysis using training, testing and validating datasets.

The rest of the paper is organized as follows. In Section 2, we discuss the experimental dataset, state of the main issue pertinent to analyzing the data for estimating the true value of the force components in three Cartesian directions using voltages read from the strain gauges mounted onto each prong of the bipolar forceps. In Section 3, we develop two nonparametric bootstrap techniques using the least squares and the weighted least squares methods, and show how the weighting approach can partially resolve some of the issues pertinent to our dataset. A multilevel modeling approach is developed in Section 4 using the Bayesian and non-Bayesian methodologies for explained estimation problems. A linear exponential (LINEX) loss function, as an example of an asymmetric penalty function, is used to obtain estimates in order to avoid underestimation of the applied forces. We also propose a weighted approach for the multilevel modeling to account for the nonhomogeneity of the read voltages from different strain gauges in the device. We discuss how to obtain the precision and confidence intervals associated with each estimate. The results of implementing the proposed approaches are reported in Section 5. Conclusions made in this research and future work are outlined in Section 6.

2 Statement of the problem and specification of employed calibration dataset

The calibration dataset is obtained by applying forces in the x, y, and z directions, where each pair of the strain gauges measures the deflection in the corresponding direction. The amount of deflection in each direction is quantified using a voltage read from the corresponding strain gauge. Specifically, the strain gauges S_1, S_2, and S_3 are designed to measure three individual voltages when a force is applied along the x-, y-, and/or z-axes, respectively. While three strain gauges are installed on each prong of the forceps, we only present the dataset of one of the prongs in this study for the sake of compactness. Figure 2 illustrates the box plots of the voltages in S_1, S_2, and S_3 when the force is applied in x, y, and z directions, respectively. When the force is applied in a certain direction, we expect to observe a substantial amount of output voltage from the corresponding strain gauge. For instance, the application of a force in the x direction should result in larger output voltages from S_1 than S_2 and S_3. Similarly, an applied force in the z direction should result in output voltages in S_3 that are significantly higher than those recorded in S_1 and S_2. The first panel of Figure 2 shows the read voltages when the force is applied in the x direction. One can see that the observed voltages in S_2 and S_3 are close to zero, while voltages read from S_1 are more significant as expected. However, when F_y is applied, an anomaly occurs as we see in the second panel of Figure 2 that the observed voltages in S_2 and S_1 are close to zero, but voltages in S_3 are remarkably significant. Nevertheless, we expected to observe voltages in S_2 be more significant than those read from S_1 and S_3. Also, from the third panel of Figure 2, it seems that when force is applied in the z direction, mounted sensors have difficulty recording anything significant. Although the behavior of the read voltages is in line with one might expect. We suspect that this problem is caused by improper placement of S_3. In other words, when a force is applied in the y direction, S_3 experiences more deflection than S_2 due to specific configuration considered to measure the voltages corresponding to the z-axis. As we observe in Figure 2, another issue pertinent to observations concerns the variability of the observed voltages in each strain gauge. Variances of the observed voltages obtained from three strain gauges are significantly different given applied forces along each axis. To address these issues, we propose methods based on the bootstrap approach with a Bayesian multilevel modeling to provide meaningful estimates of the force along the z direction.

In this paper, three groups of datasets obtained from the strain gauges S_1, S_2, and S_3 of sensorized bipolar forceps are used:(i) *Training*: in which both forces and voltages are observed, (ii) *Validation*: which is used for tuning model parameters and perform model selection (iii) and *Test*: in which we observe both voltages and forces (same as the *Training* datasets); however, the forces are assumed to be unknown. The latter group is used to predict forces by the proposed methodology for comparison purposes.[6] For obtaining the *Training* dataset, 20 trials were run using the bipolar forceps proposed and applied forces and observed voltages were recorded.[2,3] In each direction (x, y, and z), forces between 0.2 N and 2 N with increments of 0.2 N in forward and backward directions were applied to the left and right prongs of the forceps and observed voltages were recorded. In forward motion, the applied force increased from 0.2 N to 2 N, in the backward direction the force was reduces from 2 N to 0.2 N.

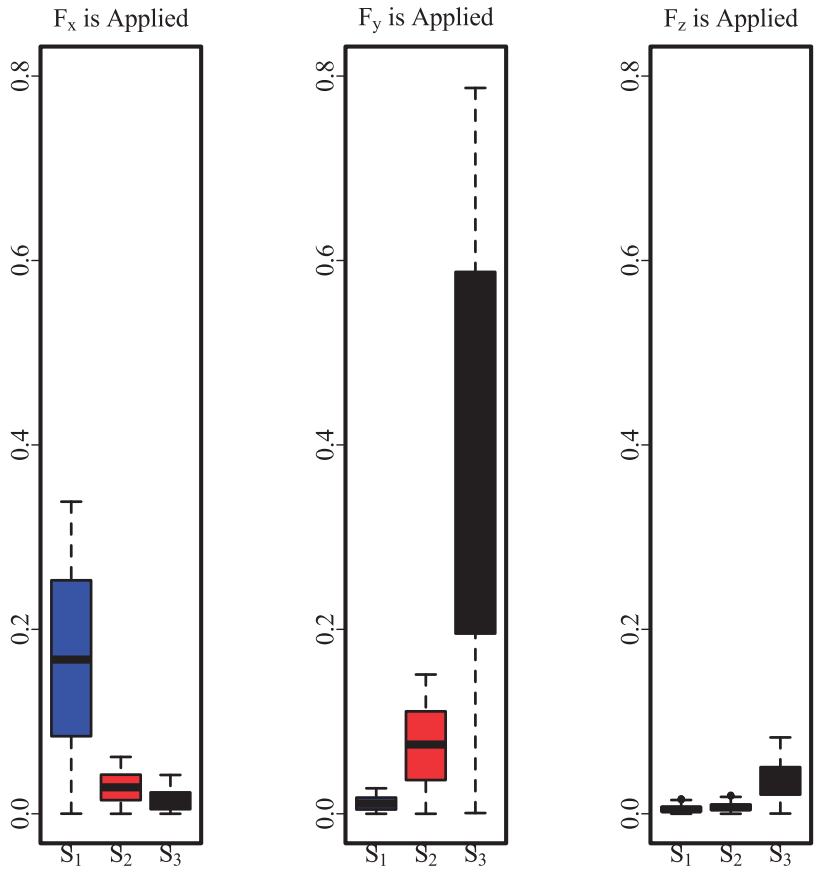

Figure 2. Observed voltages in strain gauges 1, 2, and 3 (S_1, S_2, and S_3), when force is applied in three directions. Different strain gauges are specified with different colors.

3 A bootstrap approach based on a weighted least squares method

In this section, we study a nonparametric method based on the bootstrap approach to estimate the true values of the forces using observed voltages in a 3D situation.[6] We present an algorithm based on an ordinary least squares regression modeling approach. To resolve the issue of nonhomogeneity of the read voltages and improve the bootstrap approach one can use a weighted least squares approach for model fitting.

Suppose we observe $\mathbf{y}_{l,i} = (V_{xl,i}, V_{yl,i}, V_{zl,i})^\top$, $l = 1, 2, 3$, $i = 1, \ldots, n$, where $V_{xl,i}$, $V_{yl,i}$, and $V_{zl,i}$ are voltages read at time i from strain gauges S_1, S_2, and S_3, when forces $F_{x,i}$, $F_{y,i}$, and $F_{z,i}$ are applied in x, y, and z directions, respectively. We fit the following models

$$\mathbf{y}_{l,i} = \mathbb{X}_{l,i}\mathbb{B}_l + \quad_{l,i} \tag{1}$$

where $\mathbb{B}_l = (a_l, b_l, c_l)^\top$, $\mathbb{X}_{l,i} = \text{diag}(F_{l,i}, F_{l,i}, F_{l,i})$ for all $l = 1, 2, 3$ with $F_{1,i} = F_{x,i}$, $F_{2,i} = F_{y,i}$, $F_{3,i} = F_{z,i}$ and $_{l,i} = (\epsilon_{xl,i}, \epsilon_{yl,i}, \epsilon_{zl,i})^\top$, $i = 1, \ldots, n$. We estimate the model parameters to obtain $\hat{a}_l, \hat{b}_l$, and $\hat{c}_l$, respectively and calculate fitted values $\hat{\mathbf{y}}_{l,i}$. To construct the bootstrap point and interval estimates of the applied forces, we first obtain $\hat{F}_{xl,i}$, $\hat{F}_{yl,i}$, and $\hat{F}_{zl,i}$ as solutions to the following equations

$$(\hat{V}_{xl,i}, \hat{V}_{yl,i}, \hat{V}_{zl,i})^\top = \mathbb{X}_{l,i}\hat{\mathbb{B}}_l \tag{2}$$

where $\hat{\mathbb{B}}_l = (\hat{a}_l, \hat{b}_l, \hat{c}_l)^\top$. We calculate the residuals for all l, and i, as follows

$$_{l,i} = \mathbf{y}_{l,i} - \hat{\mathbb{X}}_{l,i}\hat{\mathbb{B}}_l \tag{3}$$

with $\hat{\mathbb{X}}_{l,i} = \text{diag}(\hat{F}_{l,i}, \hat{F}_{l,i}, \hat{F}_{l,i})$ and proceed the following steps:

1. Form the rth *bootstrap* training and test datasets by resampling the residual pool $\{\boldsymbol{\epsilon}_{1,i}, \boldsymbol{\epsilon}_{2,i}, \boldsymbol{\epsilon}_{3,i}, i = 1, 2, \ldots, n\}$ as follows

$$T(r) = \{\mathbf{y}^{*r}_{l,i} = \hat{\mathbf{y}}_{l,i} + \boldsymbol{\epsilon}^{*r}_{l,i}\}^{i=1,\ldots,n}_{l=1,2,3} \tag{4}$$

and

$$V(r) = \{V^{*r}_{jl,i} = V_{jl,i} + \epsilon^{*r}_{0l,i}\}^{j=x,y,z}_{l=1,2,3} \tag{5}$$

where, $\boldsymbol{\epsilon}^{*r}_{l,i}$ and $\epsilon^{*r}_{0l,i}$ are random samples with replacement from the associated parts of the *residual pool*, each obtained independent of each other.

2. Fit models (1) to the bootstrap datasets $T(r)$ and associated forces to obtain the corresponding values of the parameters, denoted by $\hat{a}^{*r}_l$, $\hat{b}^{*r}_l$, and $\hat{c}^{*r}_l$, $l = 1, 2, 3$.
3. Calculate $\hat{F}^{*r}_{xl,i}$, $\hat{F}^{*r}_{yl,i}$, and $\hat{F}^{*r}_{zl,i}$ from (2) using the bootstrap data $V(r)$ and corresponding estimates.
4. Repeat steps 1 to 3 for $r = 1, \ldots, B$, and get B estimates.
5. Use the B estimated force values to construct point and interval estimates as described in Azimaee et al.[6]

As we show in Section 5, this method provides acceptable estimates of the forces in the x and y directions, while suffering from poor performance in the z direction. Due to the nonhomogeneity of read voltages from the strain gauges, the variability of the residuals obtained from fitting regression models in our OLS approach is also changing from one model to another. This suggests using a weighted least squares (WLS) approach with proper *weights* in order to obtain more suitable regression models to be used in our proposed bootstrap method. To this end, we use the inverse of the variances of observations read from each strain gauge to construct the necessary weights in the WLS approach to run the bootstrap calibration procedure.

4 Multilevel modeling and a Bayesian approach using the LINEX loss function

In this section, we propose a new method based on multilevel modeling. We develop necessary estimates under an asymmetric loss function to avoid underestimation of the applied forces in each direction, which is more important than overestimation in our application. Multilevel models are generalizations of the regression models, in which a set of data is structured in different categories such that their characteristics and the model's coefficients can vary within each category.

Multilevel models are also referred to as *mixed models* as they consist of fixed and random coefficients.[9] In this approach, instead of complete pooling (i.e. ignoring the effects of different group levels) or no-pooling (i.e. separate modeling for different groups), one can perform a partial pooling[8] to avoid overfitting the data and account for nonhomogeneity between different groups.

Consider the force direction as a group factor with three levels corresponding to x, y, and z, respectively. We introduce the varying slope models with no intercept[9]

$$V_{jl,i} = \beta^*_l F_{l,i} + \epsilon_{jl,i} \quad \text{with} \quad \beta^*_l = \beta_l + u_{jl} \tag{6}$$

such that $j = x, y, z$ indicates the direction of force, $F_{l,i}$ are defined as before, $l = 1, 2, 3$, refers to the index of the strain gauge that voltages $V_{jl,i}$ are read from, with $i = 1, \ldots, n_j$. For our calibration dataset $n_j = n$ for all j.

By putting together the voltages read from the stain gauge l (for all values of j and i) into a vector $\mathbf{Y}_l$ of size $3n \times 1$, one can represent (6) as

$$\mathbf{Y}_l = \beta_l \mathbf{F}_l + \mathbf{Z}_l \mathbf{u}_l + \boldsymbol{\epsilon}_l \tag{7}$$

where $\mathbf{F}_l = (\mathbf{f}_l^\top, \mathbf{f}_l^\top, \mathbf{f}_l^\top)_{3n\times 1}$ with $\mathbf{f}_l = (F_{l,1}, \cdots, F_{l,n})^\top$, β_l's are called *fixed effects*, $\mathbf{u}_l = (u_{x,l}, u_{y,l}, u_{z,l})^\top$' sare *random effects*, and

$$\mathbf{Z}_l^\top = \begin{pmatrix} F_{l,1} & \cdots & F_{l,n} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & F_{l,1} & \cdots & F_{l,n} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & F_{l,1} & \cdots & F_{l,n} \end{pmatrix}.$$

Similarly, $\boldsymbol{\epsilon}_l = (\epsilon_{xl,1}, \cdots, \epsilon_{xl,n}, \epsilon_{yl,1}, \cdots, \epsilon_{yl,n}, \epsilon_{zl,1}, \cdots, \epsilon_{zl,n})^\top$ is a $3n \times 1$ vector of error terms. Here, we have $\mathbb{E}\begin{pmatrix} \mathbf{u}_l \\ \boldsymbol{\epsilon}_l \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}$ and $\mathbb{C}ov\begin{pmatrix} \mathbf{u}_l \\ \boldsymbol{\epsilon}_l \end{pmatrix} = \begin{pmatrix} \mathbf{G}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_l \end{pmatrix}$, where $\mathbf{G}_l = \text{diag}(\sigma^2_{u_{x,l}}, \sigma^2_{u_{y,l}}, \sigma^2_{u_{x,l}})$ and $\mathbf{R}_l = \sigma^2_{\epsilon_l}\mathbb{I}_{3n\times 3n}$. Also

$$\mathbb{V}ar(\mathbf{Y}_l) = \mathbf{Z}_l\mathbf{G}_l\mathbf{Z}_l^\top + \mathbf{R}_l = \Sigma_l. \tag{8}$$

Note that working with the covariance matrix Σ_l is very convenient. However, it worth mentioning that Σ_l provides a simplified representation of the existing covariance structure in our data set and one might decide to work with more general cases where the matrix G_l is not diagonal.

4.1 Estimation of fixed and random effects coefficients

For the moment, suppose Σ_l in (8) is known. Fixed effects coefficients are normally estimated using either the maximum likelihood (ML) or generalized least squares approach. Under a multivariate normality assumption for the error terms, the log-likelihood function under the model (7) is

$$\mathbf{Z}_l\ell(\beta_l) \propto -\Big\{\log|\Sigma_l| + (\mathbf{Y}_l - \beta_l\mathbf{F}_l)^\top\Sigma_l^{-1}(\mathbf{Y}_l - \beta_l\mathbf{F}_l)\Big\}. \tag{9}$$

Maximizing (9) with respect to β_l results in

$$\hat{\beta}_l = (\mathbf{F}_l^\top\Sigma_l^{-1}\mathbf{F}_l)^{-1}\mathbf{F}_l^\top\Sigma_l^{-1}\mathbf{Y}_l.$$

Note that $\hat{\beta}_l$ performs a partial pooling by using the information obtained from the strain gauge l when forces are applied in x, y, and z directions, respectively. To estimate random effects, the best linear unbiased prediction (BLUP) approach under the squared error loss (SEL) function is often employed. Let $\delta_j(\mathbf{Y}_l)$ be an estimator for $u_{j,l}$ obtain through the following optimization problem:

$$\delta_{j,BLUP}(\mathbf{Y}_l) = \text{argmin}_{\delta_j}\,\mathbb{E}\Big[(u_{j,l} - \delta_j(\mathbf{Y}_l))^2|\mathbf{Y}_l\Big].$$

In order to derive an expression for $\delta_{j,BLUP}(\mathbf{Y}_l)$, let

$$\mathcal{H}(\delta_j(\mathbf{Y}_l)) = \mathbb{E}[u_{j,l}^2|\mathbf{Y}_l] - 2\delta_j(\mathbf{Y}_l)\mathbb{E}[u_{j,l}|\mathbf{Y}_k] + \delta_j^2(\mathbf{Y}_l).$$

which is minimized at

$$\delta_{j,BLUP}(\mathbf{Y}_l) = \mathbb{E}[u_{j,l}|\mathbf{Y}_l] \tag{10}$$

Therefore, assuming that $\mathbf{Y}_l$ and $\mathbf{u}_l$ are normally distributed, straightforward calculations show that the BLUP estimator of $\mathbf{u}_l$ under the usual SEL for a given β_l has the following form

$$\hat{\mathbf{u}}_l = \mathbb{E}(\mathbf{u}_l|\mathbf{Y}_l) = \mathbf{Z}_l^\top\mathbf{G}_l\Sigma_l^{-1}(\mathbf{Y}_l - \beta_l\mathbf{F}_l), \tag{11}$$

where, in practice, β_l is replaced by its estimate $\hat{\beta}_l$, and $\mathbf{G}_l$ and $\mathbf{R}_l$ in Σ_l are estimated as in Section 4.2.

In our calibration problem, underestimating the amount of implemented force on the brain tissue might have severe consequences such as the need for repeating a surgical task and increasing time of the operation and should be prevented. To this end, we use a LINEX loss function to penalize underestimation more than overestimation and obtain results that are more suitable for the targeted problem.[10,11] The LINEX loss function is defined as

$$L_\alpha(\hat{u}_{j,l}, u_{j,l}) = e^{\alpha(\hat{u}_{j,l}-u_{j,l})} - \alpha(\hat{u}_{j,l} - u_{j,l}) - 1$$

where the value of $\alpha > 0$ controls the penalty for underestimating the true value of $u_{j,l}$ using $\hat{u}_{j,l}$. When α approaches zero, the LINEX loss function tends to the SEL function up to a constant. In order to use the LINEX loss function, a risk unbiased predictor of $u_{j,l}$ is first found. We need to obtain $\hat{u}_{j,l}$ such that

$$\mathbb{E}[L_\alpha(\hat{u}_{j,l}, u_{j,l})] \leq \mathbb{E}[L_\alpha(\nu, u_{j,l})] \tag{12}$$

for any $nu \neq \hat{u}_{j,l}$, where expectation is taken on both $\mathbf{Y}_l$ and $u_{j,l}$.[12] When the underlying loss function is the SEL function $L(\hat{u}_{j,l}, u_{j,l}) = (\hat{u}_{j,l} - u_{j,l})^2$, the risk-unbiased predictor reduces to an unbiased predictor with $\mathbb{E}(\hat{u}_{j,l} - u_{j,l}) = 0$.

Now, the Best LINEX Unbiased Predictor of $u_{j,l}$ is obtained by minimizing $\mathbb{E}[L_\alpha(\hat{u}_{j,l}, u_{j,l})]$, with respect to $\hat{u}_{j,l}$. Using the total law of expectation and by writing $\mathbb{E}[L_\alpha(\hat{u}_{j,l}, u_{j,l})] = \mathbb{E}[\mathbb{E}[L_\alpha(\hat{u}_{j,l}, u_{j,l})]|\mathbf{Y}_{j,l}]$, we get

$$\hat{u}^\alpha_{j,l} = \frac{1}{\alpha}\log\mathbb{E}[e^{-\alpha u_{j,l}}|\mathbf{Y}_l] \tag{13}$$

Assuming that $u_{j,l}|\mathbf{Y}_l \sim N(\mu_{j,l}, \sigma^2_{j,l})$, with $\mu_{j,l} = \mathbb{E}(u_{j,l}|\mathbf{Y}_l) = \hat{u}_{j,l}$, then

$$\hat{u}^\alpha_{j,l} = \mu_{j,l} - \alpha\frac{\sigma^2_{j,l}}{2} = \hat{u}_{j,l} - \alpha\frac{\sigma^2_{j,l}}{2} \tag{14}$$

4.2 Estimation of the covariance matrices $\mathbf{G}_l$ and $\mathbf{R}_l$

The ML approach and restricted (or residual) ML (REML) strategies are proposed to estimate $\mathbf{G}_l$ and $\mathbf{R}_l$.[9] We first establish the ML approach by considering the log-likelihood function (9) as a function $\ell(\beta_l, \Sigma_l)$ of unknown parameters β_l and Σ_l.

We then substitute $\hat{\beta}_l$ in $\ell(\beta_l, \Sigma_l)$ to obtain $\ell_p(\Sigma_l)$ as follows

$$-\log|\Sigma_l| - \mathbf{Y}_l^\top\Sigma_l^{-1}(\mathbf{I} - \mathbf{F}_l(\mathbf{F}_l^\top\Sigma_l^{-1}\mathbf{F}_l)^{-1}\mathbf{F}_l^\top\Sigma_l^{-1})\mathbf{Y}_l) \tag{15}$$

where ℓ_p stands for the *profile log-likelihood* function and Σ_l is defined in (8). At this point, we can obtain the ML estimates of $\mathbf{G}_l$ and $\mathbf{R}_l$ in Σ_l, by maximizing (15) with respect to $\mathbf{R}_l$ and $\mathbf{G}_l$, respectively. For the REML estimation of $\mathbf{G}_l$ and $\mathbf{R}_l$, we maximize

$$\ell_{\text{REML}}(\Sigma_l) = \ell_p(\Sigma_l) - \frac{1}{2}\log|\mathbf{F}_l^\top\Sigma_l^{-1}\mathbf{F}_l| \tag{16}$$

with respect to $\mathbf{G}_l$ and $\mathbf{R}_l$, respectively.[9] This is a preferred approach when the sample sizes are small. However, for large samples the ML and REML estimates are very close to each other. Another technical reason for choosing the REML estimate over the ML ones is due to the fact that the influence of fixed effects' degrees of freedom is not considered in the ML approach.[13]

4.3 Bayesian approach for multilevel modeling

In order to improve our estimation in terms of accuracy, we employ the Bayesian approach for the multilevel regression problem. This allows us to incorporate prior information about the underlying problem into the

estimation process and obtain better and more accurate estimates.[9] To this end, considering the multilevel model (7), we first write the posterior distribution of $(\beta_l, \mathbf{u}_l, \mathbf{G}_l, \mathbf{R}_l)$ given the observed voltages $\mathbf{Y}_l$ as follows

$$\mathbb{P}(\beta_l, \mathbf{u}_l, \mathbf{G}_l, \mathbf{R}_l|\mathbf{Y}_l)$$

$$\propto \mathbb{P}(\mathbf{Y}_l|\beta_l, \mathbf{u}_l, \mathbf{R}_l, \mathbf{G}_l)\, \mathbb{P}(\mathbf{u}_l|\mathbf{G}_l)\mathbb{P}(\mathbf{R}_l)\, \mathbb{P}(\beta_l)\, \mathbb{P}(\mathbf{G}_l) \tag{17}$$

Suppose, $\mathbf{u}_l|\mathbf{G}_l \sim MN_3(\mathbf{0}_{3\times 1}, \mathbf{G}_l)$ and

$$\mathbf{Y}_l|\beta_l, \mathbf{u}_l, \mathbf{R}_l, \mathbf{G}_l \sim MN_{3n}(\beta_l\mathbf{F}_l + \mathbf{Z}_l\mathbf{u}_l,\ \mathbf{R}_l).$$

Therefore

$$\mathbb{P}(\mathbf{Y}_l|\beta_l, \mathbf{u}_l, \mathbf{R}_l, \mathbf{G}_l)\, \mathbb{P}(\mathbf{u}_l|\mathbf{G}_l)$$

$$\propto |\mathbf{G}_l|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\mathbf{u}_l^\top \mathbf{G}_l^{-1}\mathbf{u}_l)\right\} \times |\mathbf{R}_l|^{-\frac{1}{2}}$$

$$\times\exp\left\{-\frac{(\mathbf{Y}_l - \beta_l\mathbf{F}_l - \mathbf{Z}_l\mathbf{u}_l)^\top \mathbf{R}_l^{-1}(\mathbf{Y}_l - \beta_l\mathbf{F}_l - \mathbf{Z}_l\mathbf{u}_l)}{2}\right\} \tag{18}$$

Using (18), straightforward calculations show that $\mathbb{P}(\beta_l, \mathbf{u}_l|\mathbf{Y}_l, \mathbf{R}_l\mathbf{G}_l)$ is a multivariate normal distribution as follows

$$MN\Big((\mathbf{C}_l^\top\mathbf{C}_l + \mathbf{R}_l\mathbf{D}_l)^{-1}\mathbf{C}_l^\top\mathbf{Y}_l,\ \ \mathbf{Y}_l(\mathbf{C}_l^\top\mathbf{C}_l + \mathbf{R}_l\mathbf{D}_l)^{-1}\Big) \tag{19}$$

where, $\mathbf{C}_l \equiv \begin{pmatrix}\mathbf{F}_l & \mathbf{Z}_l\end{pmatrix}$ and $\mathbf{D}_l = \begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{G}_l^{-1}\end{pmatrix}$. Following Ruppert et al.,[9] we use an improper prior on β_l as $\mathbb{P}(\beta_l) \equiv 1$. One can also use a $N(0, \sigma^2_{\beta_l})$ distribution with very large $\sigma^2_{\beta_l}$. Furthermore, we use the following inverse gamma (IG) distributions as priors for $\sigma^2_{\epsilon_l}$ and $\{\sigma^2_{u_{j,l}}, j = x, y, z\}$, that are components of $\mathbf{R}_l$ and $\mathbf{G}_l$, respectively, and simply show them with $\mathbb{P}(\mathbf{R}_l)$ and $\mathbb{P}(\mathbf{G}_l)$

$$\begin{cases}\mathbb{P}(\mathbf{R}_l) \propto (\sigma^2_{\epsilon_l})^{-(A_{r_l}+1)}\exp\Big(-\sigma^2_{\epsilon_l}B_{r_l}\Big),\\ \mathbb{P}(\mathbf{G}_l) \propto \prod\limits_{j\in\{x,y,z\}}(\sigma^2_{u_{j,l}})^{-(A_{g_l}+1)}\exp\Big(-\sigma^2_{u_{j,l}}B_{g_l}\Big)\end{cases} \tag{20}$$

where, (A_{r_l}, B_{r_l}) and $\{(A_{g_l}, B_{g_l}), j \in \{x, y, z\}\}$ are positive *hyper-parameters*. If the hyper parameters approach to zero, then the priors for $\mathbf{R}_l$ and $\mathbf{G}_l$ would be improper and equal to $\frac{1}{\sigma^2_{\epsilon_l}}$ and $\prod_{j\in\{x,y,z\}}\frac{1}{\sigma^2_{u_{j,l}}}$, respectively. It has been recommended in Ruppert et al.,[9] to consider hyper-parameters close to zero (e.g. 0.1), to come up with non-informative, but proper priors. Based on (18)

$$\mathbb{P}(\mathbf{Y}_l|\beta_l, \mathbf{u}_l, \mathbf{R}_l, \mathbf{G}_l) \propto (\sigma^2_{\epsilon_l})^{-(\frac{3}{2}n+A_{r_l}+1)}\exp\left\{-\frac{\sigma^2_{\epsilon_l}B_{r_l}}{2}\right\}$$

$$\times\exp\left(-\frac{\sigma^2_{\epsilon_l}}{2}((\mathbf{Y}_l - \beta_l\mathbf{F}_l - \mathbf{Z}_l\mathbf{u}_l)^\top(\mathbf{Y}_l - \beta_l\mathbf{F}_l - \mathbf{Z}_l\mathbf{u}_l))\right) \tag{21}$$

Using (20) and (21), it is straightforward to show that

$$[\mathbf{R}_l|\mathbf{Y}_l,\beta_l,\mathbf{u}_l,\mathbf{G}_l]\sim IG\left(A_{r_l}+\frac{n}{2}\right)$$

$$(\mathbf{Y}_l-\beta_l\mathbf{F}_l-\mathbf{Z}_l\mathbf{u}_l)^\top(\mathbf{Y}_l-\beta_l\mathbf{F}_l-\mathbf{Z}_l\mathbf{u}_l)+B_{r_l}) \tag{22}$$

Similar to (21) for $\mathbf{G}_l$ one can show that $[\mathbf{G}_l|\mathbf{Y}_l,\beta_l,\mathbf{u}_l,\mathbf{R}_l]$ follows

$$IG\left(A_{g_l}+\frac{3}{2},B_{g_l}+\frac{1}{2}\mathbf{u}_l^\top\mathbf{u}_l\right) \tag{23}$$

Finally, in order to make statistical inference, we use the Markov Chain Monte Carlo (MCMC) approach to sample from the posterior distribution based on the following algorithm[9]

1. Generate initial values $\Psi^{(0)}=(\beta_l^{(0)},\mathbf{u}_l^{(0)},\mathbf{R}_l^{(0)},\mathbf{G}_l^{(0)})$ using the ML or REML approaches as explained in Sections 4.1 and 4.2.
2. For $t=0,1,\ldots,M$, where M is the number of iterations, repeat the following steps:

- Sample $(\beta_l^{(t+1)},\mathbf{u}_l^{(t+1)})$ from $\mathbb{P}(\beta_l,\mathbf{u}_l|\mathbf{Y}_l^{(t)},\mathbf{R}_l^{(t)}\,\mathbf{G}_l^{(t)})$ in (19).
- Sample $\mathbf{G}_l^{(t+1)}$ from (23) given $\Psi^{(t)}$.
- Sample $\mathbf{R}_l^{(t+1)}$ from (22) given $\Psi^{(t)}$.

For the comparison purpose, we obtain the estimates under the SEL and LINEX loss functions. Through the MCMC process, we obtain M estimates of β_l and $\mathbf{u}_l$, where under the SEL, we consider the mean value of these samples $\hat{u}$ and $\hat{\beta}$ as follows

$$\hat{\mathbf{u}}_l=\frac{1}{M}\sum_{t=1}^{M}\mathbf{u}_l^{(t)}\quad\text{and}\quad\hat{\beta}_l=\frac{1}{M}\sum_{t=1}^{M}\beta_l^{(t)}. \tag{24}$$

Under the LINEX loss function these estimates are obtained by averaging estimates obtained via (14), $\hat{u}$ and $\hat{\beta}$, respectively.

5 Results

5.1 Results obtained from bootstrapping

Table 1 represents the estimated bootstrap confidence bounds in each direction for the amount of forces between 0.2 N and 2 N with an increment of 0.2 N. Under the LS approach the bootstrap confidence intervals are too wide, and the estimated forces are often very different from the true force values. Figure 3 shows the observed voltages from each strain gauge where different colors are used to show different directions. It also shows the residuals obtained by fitting LS regression models in (1). One can see that the residuals are not homogeneous and those pertinent to S_3 are relatively large. To adjust for nonhomogeneity of the error terms, we used the bootstrap method with a WLS approach based on the following weights[7]

$$W_l=(\mathbb{V}ar^{-1}(V_{xl}),\mathbb{V}ar^{-1}(V_{yl}),\mathbb{V}ar^{-1}(V_{zl})),\quad l=1,2,3 \tag{25}$$

These weights are associated with the variance of the voltages read from S_l in (1), $l=1,2,3$. Results for the bootstrap method with WLS regression models are presented in Table 2. We observe that estimated confidence intervals are significantly improved compared to Table 1. For most cases, the true amount of forces in the x and y directions are within the estimated intervals. However, estimated confidence intervals for forces in the z direction are still not accurate enough, and the length of the intervals is considerably large.

Table 1. Bootstrap confidence intervals for forces applied to the right forceps tip, in the *x*, *y*, and *z* directions using the LS regression approach.

True force	F_x C.I	F_y C.I	F_z C.I
0.2	(0.000, 0.893)	(0.000, 2.000)	(0.000, 2.000)
0.4	(0.275, 2.000)	(0.000, 2.000)	(0.000, 2.000)
0.6	(0.447, 2.000)	(0.777, 2.000)	(0.000, 2.000)
0.8	(0.994, 2.000)	(0.851, 2.000)	(0.000, 2.000)
1.0	(0.000, 1.105)	(1.370, 2.000)	(0.000, 2.000)
1.2	(0.314, 1.691)	(0.000, 2.000)	(0.000, 2.000)
1.4	(0.890, 2.000)	(0.442, 2.000)	(0.000, 2.000)
1.6	(0.728, 2.000)	(0.000, 2.000)	(0.000, 2.000)
1.8	(1.263, 2.000)	(0.974, 2.000)	(0.000, 2.000)
2.0	(1.269, 2.000)	(1.324, 2.000)	(0.000, 2.000)

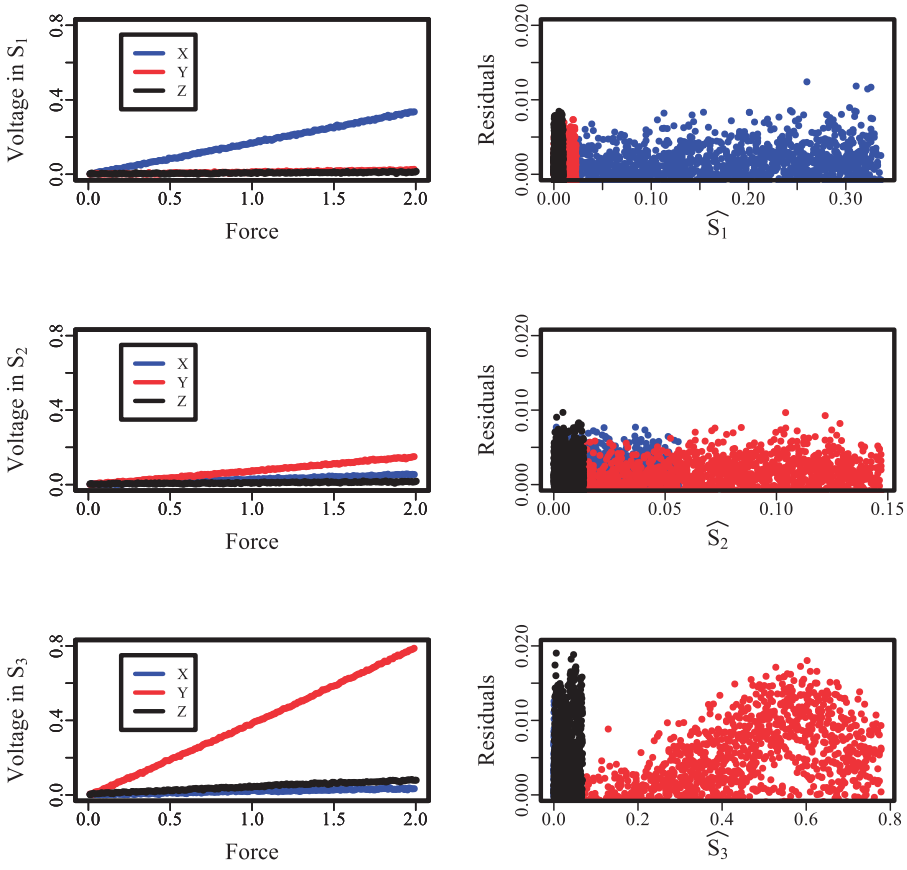

Figure 3. In the left panel, observed voltages from each strain gauge are presented versus the amount of force in three directions. In the right panel, residuals obtained from fitting models in each strain gauge are illustrated versus the fitted values of voltages for all three directions. Different directions are specified with different colors.

5.2 Results obtained from multilevel modeling

For the multilevel modeling approach we used **R** and the built-in function *lmer*, within the *lme4* package.[14]

In theory, we should not consider intercepts in the underlying models, since voltages are not observed when no forces are applied. However, we violate this assumption and consider models with intercept, which provides better results. Also, we performed our study under both SEL and LINEX loss functions. For the LINEX loss function, we use $\alpha \in \{0.1, 0.5, 1\}$ to provide different measures of penalty for overestimation compared with underestimation.

Table 2. Confidence intervals for forces applied to the right forceps tip, in the *x*, *y*, and *z* directions using the WLS regression approach.

True force	F_x C.I	F_y C.I	F_z C.I
0.2	(0.156, 0.199)	(0.169, 0.264)	(0.290, 0.766)
0.4	(0.394, 0.438)	(0.347, 0.442)	(0.378, 0.867)
0.6	(0.580, 0.625)	(0.589, 0.682)	(0.657, 1.143)
0.8	(0.760, 0.805)	(0.791, 0.886)	(1.050, 1.530)
1.0	(0.975, 1.020)	(0.998, 1.090)	(1.153, 1.622)
1.2	(1.199, 1.243)	(1.216, 1.310)	(1.636, 2.000)
1.4	(1.390, 1.436)	(1.389, 1.483)	(1.739, 2.000)
1.6	(1.586, 1.630)	(1.566, 1.659)	(1.859, 2.000)
1.8	(1.800, 1.846)	(1.801, 1.897)	(2.000, 2.000)
2.0	(1.976, 2.000)	(2.000, 2.000)	(2.000, 2.000)

Table 3. Six different proposed weights to use in multilevel modeling approach.

W_1	W_2	W_3	W_4	W_5	W_6
1	1	1	$V_{y,1}$	$Y_{y,1}$	1
1	1	1	$V_{y,2}$	$V_{y,2}$	1
$\frac{1}{\mathbb{V}ar(V_{z,3})}$	$\frac{1}{\mathbb{V}ar(V_{y,2})}$	$\frac{1}{V_{y,2}}$	$V_{y,3}$	$V_{x,1}$	$\frac{1}{\text{mean}(V_{z,3})}$

According to our numerical studies (not presented here), we observed that under the LINEX loss function, estimated forces in the x and y directions are close to their true values, but estimation in the z direction is not accurate enough irrespective of the values of α. The details of such numerical studies can be found in Chapter 4 of Azimaee.[15] This is due to mechanical issues associated with adding the third pair of strain gauges as discussed in Section 2. To address this, we decided to apply weights in our multilevel modeling approach. Since the problem was only in estimating the forces along z direction, we applied weight only for $l = 3$ in model (7) to account for the problem in the voltages read from S_l. As shown in Table 3, we considered different weights associated with data obtained from S_3 when forces are applied in the x, y, and z directions.

Our comprehensive numerical studies in Azimaee[15] show that models with intercept provide better force estimation compared to those without intercept. As such, we only present the results for models with an intercept from here on. Table 4 shows estimated forces in the z direction using different weighs under the SEL function for models with and without intercept. Under the LINEX loss function the best results are obtained when $\alpha = 0.5$ and we use W_1 in the model (see Table 5). Also, the results under the SEL and LINEX loss functions are very similar.

In order to provide interval estimation, we employed a bootstrap technique with multilevel models. Once again, we considered both models with and without an intercept for obtaining the results. Bootstrap results are obtained from models with considering weights (WLS). The bootstrap procedure is the same as in Section 3, except that instead of linear regression models, we fitted multilevel models as presented in (7). Based on the results presented in Tables 4 and 5, we selected W_3 to be used in the model for estimating forces in z direction under SEL and W_1 under LINEX loss in the bootstrapping procedure. We also considered $\alpha = 0.1$, since the results under this value were considerably better than other values. In most cases, true values of the applied forces are within the estimated intervals and the bias is very low. However, there are some differences between the results obtained for applied forces in the z direction from models with intercept and models without intercept. While there are wider confidence bounds for models with intercepts, the bias and RMSE values, are lower for these models and most of the time the true value of the force is within the estimated interval.

5.3 Results obtained using the Bayesian approach

In this section, we report predicted values of the applied forces along x, y, and z directions that are obtained using the Bayesian approach under both SEL and LINEX loss functions. To obtain the results, we used *lmer-stan* function of "*rstanarm*" package, in **R** programming language.[16]

Table 4. Estimated forces in the z direction, considering six different weights under the SEL loss function using models with or without intercepts.

	$\hat{F}_z$ in models without intercept						$\hat{F}_z$ in models with intercept					
True F_z	W_1	W_2	W_3	W_4	W_5	W_6	W_1	W_2	W_3	W_4	W_5	W_6
0.2	0.30	0.30	0.29	0.31	0.31	0.30	0.15	0.15	0.21	0.16	0.19	0.17
0.4	0.52	0.52	0.51	0.53	0.53	0.52	0.41	0.41	0.44	0.42	0.44	0.42
0.6	0.70	0.70	0.67	0.71	0.71	0.70	0.60	0.60	0.63	0.61	0.63	0.62
0.8	0.85	0.85	0.82	0.87	0.87	0.85	0.78	0.78	0.80	0.79	0.80	0.79
1.0	1.09	1.09	1.05	1.11	1.11	1.09	1.06	1.06	1.06	1.06	1.07	1.06
1.2	1.27	1.27	1.22	1.29	1.29	1.27	1.25	1.25	1.24	1.26	1.26	1.26
1.4	1.41	1.41	1.36	1.43	1.43	1.41	1.42	1.42	1.40	1.42	1.42	1.42
1.6	1.56	1.56	1.50	1.58	1.58	1.56	1.58	1.58	1.55	1.58	1.58	1.58
1.8	1.70	1.70	1.64	1.73	1.73	1.70	1.75	1.75	1.71	1.75	1.75	1.74
2.0	1.94	1.94	1.87	1.97	1.97	1.94	2.00	2.00	1.97	2.00	2.00	2.00

Table 5. Estimated forces in z direction, considering six different weights under the LINEX loss using models with intercept using the multilevel models and the Bayesian approach.

	$\alpha_1 = 0.1$						$\alpha_2 = 0.5$						$\alpha_3 = 1$					
True F_z	W_1	W_2	W_3	W_4	W_5	W_6	W_1	W_2	W_3	W_4	W_5	W_6	W_1	W_2	W_3	W_4	W_5	W_6
Using multilevel models																		
0.2	0.16	0.18	0.21	0.16	0.18	0.21	0.21	0.24	0.27	0.17	0.19	0.22	0.19	0.22	0.25	0.18	0.20	0.23
0.4	0.42	0.47	0.55	0.42	0.47	0.55	0.46	0.51	0.59	0.43	0.48	0.56	0.45	0.50	0.59	0.43	0.48	0.57
0.6	0.62	0.70	0.82	0.62	0.70	0.82	0.65	0.72	0.84	0.63	0.71	0.83	0.65	0.72	0.85	0.63	0.71	0.83
0.8	0.80	0.90	1.06	0.80	0.90	1.06	0.82	0.91	1.06	0.81	0.91	1.07	0.82	0.92	1.08	0.81	0.91	1.07
1.0	1.08	1.22	1.43	1.08	1.22	1.43	1.08	1.21	1.41	1.09	1.22	1.44	1.10	1.23	1.44	1.09	1.22	1.43
1.2	1.29	1.44	1.70	1.29	1.44	1.70	1.27	1.42	1.65	1.29	1.45	1.70	1.30	1.45	1.70	1.29	1.44	1.70
1.4	1.46	1.63	1.93	1.46	1.63	1.93	1.43	1.60	1.86	1.46	1.64	1.93	1.46	1.63	1.92	1.46	1.63	1.92
1.6	1.63	1.83	2.15	1.63	1.82	2.15	1.59	1.78	2.07	1.63	1.82	2.15	1.63	1.82	2.14	1.62	1.82	2.14
1.8	1.80	2.02	2.38	1.80	2.01	2.38	1.75	1.95	2.28	1.80	2.01	2.37	1.79	2.00	2.35	1.79	2.00	2.36
2.0	2.08	2.33	2.75	2.08	2.33	2.75	2.02	2.25	2.62	2.08	2.33	2.74	2.07	2.31	2.71	2.07	2.32	2.72
Using the Bayesian approach																		
0.2	0.157	0.184	0.232	0.155	0.164	0.178	0.208	0.214	0.224	0.169	0.180	0.195	0.193	0.209	0.233	0.172	0.178	0.186
0.4	0.420	0.490	0.619	0.412	0.436	0.472	0.448	0.463	0.483	0.424	0.450	0.488	0.445	0.481	0.536	0.424	0.438	0.458
0.6	0.625	0.729	0.922	0.613	0.649	0.702	0.636	0.657	0.686	0.623	0.662	0.717	0.642	0.694	0.774	0.620	0.642	0.671
0.8	0.808	0.943	1.191	0.791	0.839	0.906	0.804	0.830	0.866	0.800	0.850	0.921	0.817	0.884	0.985	0.795	0.823	0.860
1.0	1.093	1.275	1.612	1.070	1.134	1.226	1.065	1.100	1.147	1.077	1.144	1.239	1.090	1.180	1.314	1.069	1.106	1.155
1.2	1.298	1.515	1.915	1.271	1.347	1.456	1.253	1.295	1.350	1.276	1.355	1.469	1.287	1.393	1.552	1.266	1.309	1.368
1.4	1.469	1.715	2.167	1.439	1.525	1.648	1.410	1.457	1.519	1.442	1.532	1.660	1.452	1.570	1.750	1.430	1.479	1.546
1.6	1.641	1.914	2.419	1.606	1.702	1.840	1.567	1.619	1.688	1.609	1.708	1.851	1.616	1.748	1.947	1.594	1.649	1.723
1.8	1.812	2.114	2.672	1.774	1.880	2.031	1.724	1.781	1.857	1.775	1.884	2.042	1.780	1.926	2.145	1.758	1.818	1.900
2.0	2.097	2.447	3.092	2.053	2.175	2.351	1.985	2.051	2.139	2.051	2.178	2.360	2.053	2.221	2.475	2.031	2.101	2.196

Once again, we considered intercepts in the model and applied weights (as described in Section 3) in the model (WLS). Results are reported in Table 5 for models with an intercept under the LINEX loss function. One can observe that estimated forces using models with intercepts are close to the true values of forces. Among all weights that have been employed, W_4 provides more accurate results and most predictions are very close to the true amount of applied forces for models under both SEL and LINEX loss functions. For instance, there is no bias in estimating the true force of 2 N, under the SEL when W_4 is applied. Another point to consider is the robustness of the algorithm with respect to the choice of different priors for (β, u). To check this, we obtained the force prediction in x, y, and z directions under the SEL function considering noninformative, normal, and Student's t priors. As one can observe in Table 6, there is not any significant difference between estimates under different priors.

Table 6. Point estimates of applied forces in the *x*, *y*, and *z* directions using the Bayesian approach in multilevel modeling with different priors for models with intercepts.

Prior	Noninformative			Normal			Student's *t*		
True force	$\hat{F}_x$	$\hat{F}_y$	$\hat{F}_z$	$\hat{F}_x$	$\hat{F}_y$	$\hat{F}_z$	$\hat{F}_x$	$\hat{F}_y$	$\hat{F}_z$
0.2	0.176	0.183	0.324	0.176	0.183	0.321	0.176	0.183	0.320
0.4	0.411	0.444	0.543	0.411	0.444	0.539	0.411	0.444	0.537
0.6	0.577	0.583	0.715	0.578	0.583	0.710	0.577	0.584	0.707
0.8	0.801	0.731	0.867	0.801	0.730	0.862	0.801	0.731	0.858
1.0	0.982	0.967	1.106	0.982	0.966	1.099	0.982	0.967	1.094
1.2	1.222	1.224	1.283	1.224	1.225	1.340	1.225	1.226	1.419
1.4	1.418	1.397	1.420	1.418	1.396	1.412	1.417	1.398	1.405
1.6	1.581	1.650	1.563	1.581	1.649	1.554	1.580	1.651	1.546
1.8	1.824	1.852	1.706	1.825	1.851	1.696	1.824	1.853	1.688
2.0	2.018	2.059	1.945	2.019	2.058	1.933	2.018	2.060	1.923

Table 7. Point and interval estimates of applied forces in the *x*, *y*, and *z* directions using the bootstrap method and the Bayesian approach for multilevel modeling under the SEL and LINEX loss functions.

	Force in *x* direction				Force in *y* direction				Force in *z* direction			
True force	C.I	$\hat{F}$	Bias	RMSE	C.I	$\hat{F}$	Bias	RMSE	C.I	$\hat{F}$	Bias	RMSE
Estimation under the SEL function												
0.2	(0.162, 0.203)	0.175	−0.025	0.030	(0.161, 0.257)	0.185	−0.015	0.038	(0.074, 0.593)	0.148	−0.052	0.335
0.4	(0.397, 0.437)	0.409	0.009	0.018	(0.423, 0.518)	0.447	0.047	0.059	(0.565, 1.116)	0.666	0.266	0.424
0.6	(0.563, 0.603)	0.576	−0.024	0.029	(0.562, 0.658)	0.586	−0.014	0.038	(0.575, 1.127)	0.677	0.077	0.339
0.8	(0.786, 0.827)	0.799	−0.001	0.016	(0.709, 0.805)	0.733	−0.067	0.075	(0.666, 1.225)	0.774	−0.026	0.332
1.0	(0.967, 1.008)	0.981	−0.019	0.025	(0.945, 1.042)	0.969	−0.031	0.047	(1.037, 1.584)	1.129	0.129	0.356
1.2	(1.207, 1.249)	1.221	0.021	0.026	(1.201, 1.299)	1.227	0.027	0.044	(1.105, 1.649)	1.194	−0.006	0.332
1.4	(1.403, 1.445)	1.417	0.017	0.023	(1.374, 1.473)	1.399	−0.001	0.036	(1.390, 1.933)	1.474	0.074	0.341
1.6	(1.566, 1.608)	1.579	−0.021	0.026	(1.626, 1.726)	1.652	0.052	0.063	(1.546, 2.096)	1.636	0.036	0.336
1.8	(1.810, 1.852)	1.823	0.023	0.028	(1.827, 1.929)	1.855	0.055	0.065	(1.698, 2.248)	1.787	−0.013	0.336
2.0	(2.004, 2.046)	2.017	0.017	0.024	(2.033, 2.136)	2.061	0.061	0.071	(1.834, 2.390)	1.927	−0.073	0.344
Estimation under the LINEX loss function with $\alpha = 1$ and the weight function W_4.												
0.2	(0.176, 0.180)	0.176	−0.024	0.028	(0.181, 0.190)	0.183	−0.017	0.045	(0.271, 0.406)	0.134	−0.066	0.396
0.4	(0.412, 0.415)	0.409	0.009	0.016	(0.443, 0.451)	0.444	0.044	0.061	(0.910, 1.034)	0.658	0.258	0.468
0.6	(0.580, 0.583)	0.576	−0.024	0.028	(0.582, 0.590)	0.583	−0.017	0.045	(0.924, 1.049)	0.669	0.069	0.397
0.8	(0.804, 0.807)	0.800	−0.000	0.014	(0.731, 0.737)	0.730	−0.070	0.081	(1.041, 1.176)	0.767	−0.033	0.393
1.0	(0.987, 0.989)	0.984	−0.016	0.022	(0.967, 0.974)	0.966	−0.034	0.054	(1.457, 1.643)	1.126	0.126	0.413
1.2	(1.228, 1.231)	1.222	0.022	0.026	(1.224, 1.232)	1.224	0.024	0.048	(1.536, 1.731)	1.192	−0.008	0.394
1.4	(1.424, 1.428)	1.418	0.018	0.023	(1.396, 1.406)	1.396	–0.004	0.042	(1.857, 2.120)	1.475	0.075	0.403
1.6	(1.588, 1.592)	1.582	−0.018	0.023	(1.649, 1.661)	1.648	0.048	0.063	(2.044, 2.339)	1.639	0.039	0.399
1.8	(1.833, 1.837)	1.826	0.026	0.029	(1.851, 1.865)	1.849	0.049	0.064	(2.218, 2.544)	1.791	−0.009	0.398
2.0	(2.028, 2.033)	2.020	0.020	0.025	(2.057, 2.074)	2.054	0.054	0.068	(2.383, 2.734)	1.933	−0.067	0.405

RMSE: root mean square error.

We also employed the proposed bootstrap method in Section 3 for the Bayesian approach. Results of bootstrapping under the SEL approach and the LINEX loss function are presented in Table 7 when we use the weight function W_4 and $\alpha = 1$. Under the SEL function confidence bounds that we obtained using models with intercept for F_z are more precise and almost always contain the true values of the forces, nevertheless, the bounds are relatively wide. Point estimates of applied forces in the z direction are much better than other estimates and the bias is negligible under the SEL function. As can be seen, under the LINEX loss function, estimated confidence intervals for F_x and F_y are very narrow, and often do not contain the true values of the force. Although the calculated bounds for F_z are also narrower, true values of force are not included within the bounds. The advantage of results obtained under the LINEX loss is that the point estimates with considering intercept in the models

are relatively accurate. In Table 7, we have used the bias and RMSE as the usual performance measures under to compare our estimates. One can also calculate the bias and risk associated with estimates under the LINEX loss function.[17]

6 Conclusions and future work

We proposed several statistical methodologies to provide accurate point and interval estimates of force components exerted to the brain tissue during the performance of surgical tasks. Three small strain gauges were mounted onto each prong of a conventional surgical bipolar forceps, and the voltage signals of each strain gauge installed on a calibration station were used to estimate the value of force components applied. To calculate the estimates of the interaction forces, we developed a bootstrap technique that was later used in conjunction with weighted least-squares linear regression and multilevel models to estimate the unknown force values using observed voltages. The nonparametric bootstrap technique provides accurate point and interval force estimation along the x and y directions and fails to accurately estimate the force in the z direction due to particular mechanical design considered to place the strain gauge for measuring the force along the z-axis. We introduced different weights to the model building process to come up with more accurate estimates. After applying these weights, bootstrap results were improved as well, but still, results were not satisfactory for the z direction.

Since the calibration data was readily structured based on the direction in which the force was exerted to the brain tissue (in x, y, and z directions), we developed a proper multilevel model for our calibration problem. The force estimates obtained from the multilevel modeling yield better estimates when compared with those under the WLS approach. The estimates improved further as weights were introduced into the model. Furthermore, we enhanced our results by using a Bayesian approach to estimate the coefficients of the underlying multilevel models. Results obtained under both the SEL and the LINEX loss functions to avoid overestimating the applied forces. We also employed our proposed bootstrap methodology in conjunction with multilevel models to provide interval estimates and enhance the accuracy of our point estimation. The overall results suggest that estimations obtained using multilevel models are more accurate; however, the Bayesian approach provides narrower confidence bounds with smaller bias and RMSEs. Another important observation is that using an intercept in the model helps obtain more accurate results. An interesting direction for future work is to consider measurement error in the model, specifically, multilevel models.[18] This seems to be a proper direction as there are obviously measurement errors in the read voltages from the strain gauges. Another interesting research direction involves robust Bayesian analysis to study the sensitivity of our results with respect to the choice of prior distributions in the underlying Bayesian multilevel modeling approach.

Acknowledgment

The authors gratefully acknowledge the contributions of Project neuroArm at the University of Calgary for sharing the data set for validating the statistical tools employed in this work. These set of data were previously used in part for calibration of surgical tools under the nonparametric Bootstrap technique in 2D.[6] We would like to thank reviewers for their constructive comments, which improved the quality of this work.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Mohammad Jafari Jozani acknowledges the partial support of NSERC.

ORCID iD

Mohammad Jafari Jozani https://orcid.org/0000-0002-2147-6337

References

1. Maddahi Y, Huang J, Huang J et al. Real-time measurement of tool-tissue interaction forces in neurosurgery: quantification and analysis. In: 2016 IEEE international *conference on advanced intelligent mechatronics (AIM)*, 2016, pp. 1405–1410. New York: IEEE.

2. Zareinia K, Maddahi Y, Gan LS et al. A force-sensing bipolar forceps to quantify tool–tissue interaction forces in microsurgery. *IEEE/ASME Trans Mechatron* 2016; **21**(5): 2365–2377.
3. Gan LS, Zareinia K, Lama S et al. Quantification of forces during a neurosurgical procedure: a pilot study. *World Neurosurg* 2015; **84**: 537–548.
4. Gorbushin A and Bolshakova A. Unsteady axial force measurement by the strain gauge balance. *Measurement* 2020; **152**: 107381.
5. Sugiyama T, Lama S, Gan LS, et al. Forces of tool-tissue interaction to assess surgical skill level. *JAMA Surg* 2018; **153**: 234–242.
6. Azimaee P, Jafari Jozani M, Maddahi Y et al. Nonparametric bootstrap technique for calibrating surgical smartforceps: theory and application. *Expert Rev Med Dev* 2017; **14**(10): 833–843.
7. Draper NR and Smith H. *Applied regression analysis*. New York: John Wiley & Sons, 2014.
8. Gelman A and Hill J. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press, 2006.
9. Ruppert D, Wand MP and Carroll RJ. *Semiparametric regression*. Cambridge: Cambridge University Press, 2003.
10. Arashi M and Tabatabaey SMM. Estimation of the location parameter under LINEX loss function: multivariate case. *Metrika* 2010; **72**: 51–57.
11. Torehzade S and Arashi M. A note on shrinkage wavelet estimation in Bayesian analysis. *Stat Prob Lett* 2014; **84**: 231–234.
12. Chattopadhyay S, Chaturvedi A and Sengupta RN. LINEX loss function and its statistical applications-a review. *Decision* 1999; **26**: 51.
13. Searle SR, Casella G and McCulloch CE. *Variance components*. **vol. 391**. New York: John Wiley & Sons, 2009.
14. De Boeck P, Bakker M, Zwitser R, et al. The estimation of item response models with the lmer function from the lme4 package in *r*. *J Stat Softw* 2011; **39**: 1–28.
15. Azimaee P. *Calibratings surgical SmartForceps* ™ *using bootstrap and multilevel modeling techniques*. Master's Thesis, University of Manitoba, Canada, 2017.
16. Gabry J and Goodrich B. rstanarm: Bayesian applied regression modeling via stan. r package version 2.10. 0, 2016.
17. Parsian A and Sanjari Farsipour N. Estimation of the mean of the selected population under asymmetric loss function. *Metrika* 1999; **50**.
18. Goldstein H. Measurement errors in multilevel models. In: Shewhart WA and Wilks SS (eds) *Multilevel statistical models*. US: John Wiley & Sons, Ltd, 2010, pp. 267–284.