Research paper

Computational load estimation of the femur

Gianni Campoli\textsuperscript{a,1}, Harrie Weinans\textsuperscript{a,b}, Amir Abbas Zadpoor\textsuperscript{a,*1}

\textsuperscript{a} Department of Biomechanical Engineering, Faculty of Mechanical, Maritime, and Materials Engineering, Delft University of Technology (TU Delft), Mekelweg 2, Delft 2628 CD, The Netherlands
\textsuperscript{b} Department of Orthopaedics, Erasmus University Medical Center, Rotterdam, The Netherlands

\textbf{ARTICLE INFO}

Article history:
Received 30 August 2011
Received in revised form 1 February 2012
Accepted 19 February 2012
Published online 27 February 2012

Keywords:
Load prediction
Bone remodeling
Musculoskeletal loads
Modeling of bone tissue adaptation
Finite element method
Femur
Artificial neural networks
Wavelets

\textbf{ABSTRACT}

The density distribution and, thus, mechanical properties of long bones such as the femur are dependent on their loading. Many bone tissue adaptation theories are proposed to describe the density distribution that results from a given set of loading parameters. It is relatively easy to measure the density distribution of long bones, for example, using Computed Tomography (CT). However, there is no easy non-invasive method for in-vivo measurement of musculoskeletal loads. It is therefore interesting to investigate whether or not it is possible to predict the musculoskeletal loads that have resulted in a certain measured density distribution using bone tissue adaptation models. An inverse problem has to be solved for that purpose. In this paper, we use Artificial Neural Networks (ANNs) to solve the associated inverse problem and estimate the loading parameters that have resulted in the CT-measured three-dimensional density distribution of a proximal femur. An ANN is trained using a dataset generated by solving the forward tissue adaptation model for a large number of loading parameters. Before training the ANN with the generated training dataset, a Gaussian noise component is added to the density distribution. This improves the robustness of the trained ANN against deviations of the measured density distribution from the predictions of the forward bone tissue adaptation model. It is shown that the proposed technique is capable of predicting loading parameters that result in a density distribution close to the measured density distribution.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

According to Wolff’s law, bones adapt to the loads they experience. When loads increase, for example, when a person becomes physically more active, more bone is added to strengthen the tissue (Borer, 2005; Maimoun and Sultan, 2011). When loads decrease, for example, in prolonged bed rest (Zerwekh et al., 1998) or in prolonged exposure to microgravity conditions (Cowin, 1998; Vico et al., 1988), some bone is removed and the mass of the tissue decreases. Bone tissue adaptation may happen locally, for example, due to redistribution of mechanical stimulus caused by implantation of a prosthesis (Engh et al., 1992; Huiskes et al., 1987). A number of mathematical models are developed to predict the adaptation of bone tissue caused by a certain loading pattern. These models are based on certain assumptions about the
way bone tissue adaptation works. For example, a large number of models, e.g. (Cowin and Hegedus, 1976; Doblare and Garcia, 2001; Huiskes et al., 1987), assume that the bone tissue adaptation mechanism tries to maintain some sort of homeostatic state of mechanical stimulus (stress, strain, etc.) by forming or resorbing bone. Other models are based on a global optimality criterion (Fernandes et al., 1999) or on damage accumulation/repair postulates (Prendergast and Taylor, 1994). Mathematical models of bone tissue adaptation are normally implemented in finite element (FE) programs to enable numerical solution of various problems related to bone tissue adaptation. A large number of studies have used these computational models for such important applications as implant design, e.g. (Friedman et al., 1992; Huiskes et al., 1987, 1989; Lin et al., 2010; Orr et al., 1990).

If it is possible to start from the loads and predict the resulting morphology of bone, it may be possible to start from a given morphology of bone and try to predict the loading history that has resulted in the given morphology. The possibility of solving this inverse problem opens up many opportunities. For example, one can estimate the musculoskeletal loads and the characteristics of the daily activities that have resulted in the measured density distribution (Bona et al., 2006a) or can distinguish between different modes of human/animal locomotion (Bona, 2003). Moreover, the method can be used for estimating the musculoskeletal loads when direct measurement is not possible such as in the case of fossil bones (Christen et al., 2011). The other application would be validation of the muscle and joint reaction forces that are predicted by large-scale musculoskeletal models, as there is no non-invasive method for in-vivo measurement of such loads.

Despite the importance of the inverse problem, there has been only limited research on this topic so far, perhaps due to the complexity of the techniques that are required for solving the inverse problem. The most notable works are a series of studies by Fischer and his collaborators (Bona, 2003; Bona et al., 2006b; Fischer et al., 1999, 2003, 1995, 1996, 1998). Research on this topic seems to have been recently revived (Christen et al., 2011; Zadpoor et al., submitted for publication) thanks to the availability of improved computational power and some methodological developments that have made solving the inverse problem easier and more realistic.

In a recent study (Zadpoor et al., submitted for publication), we proposed a new method based on Artificial Neural Networks (ANNs) for prediction of the loading history based on the measured morphology of trabecular bone. In the proposed methodology, the ANN works as a nonlinear mapping from the space of density distribution to the space of loading parameters. In the current study, we apply our recently proposed methodology to predict the likely loading experienced by a proximal femur. A number of novelties separate the current work from our previous study. First, the problem considered in the previous study was at the tissue level (trabecular bone) whereas organ level (proximal femur) load prediction is considered in the current study. Second, as opposed to the previous study, the problem considered in the present study is three-dimensional. Third, in the current study, the load prediction is based on the measured density distribution of a cadaveric femur. Load prediction was carried out using simulated density distributions in the previous study. Finally, a new method based on wavelet decomposition is proposed in the current study for dealing with the large number of inputs that are required for introducing the density distribution of a discretized proximal geometry to the ANN.

2. Methodology

The modeling scheme that is proposed for neural network prediction of the load from density distribution consists of three parts. The first three subsections of the current section cover those three parts. The first part is an FE model that is capable of calculating the stress and strain distributions that result from a certain loading condition (Section 2.1). The second part is a forward tissue adaptation model that can predict the tissue adaptation caused by a certain stress and strain distribution (Section 2.2). The third part is the ANN that is trained using the results of the forward tissue adaptation model (Section 2.3). The final subsection describes the methodology that is used for studying the convergence of the obtained solutions.

2.1. Description of the FE model

Three three-dimensional meshes of a cadaveric femur were obtained from the website of the VAKHUM project (http://www.ulb.ac.be/project/vakhum/). A complete description of the methodology used for generating the FE meshes is presented in the documentation of the VAKHUM project. Therefore, only a brief description is presented here. According to the documentation of the VAKHUM project, a dual-slice beam CT-scanner (Elscint Spiral Twin Flash) was used for scanning a cadaveric femur in vitro. The obtained images were segmented semi-automatically and were subsequently used to generate iso-surface 3D models using an algorithm that was based on the Marching Cubes method (Lorensen and Cline, 1987). From the 3D bone models, an advanced implementation of the original hexahedral mesh generator proposed by Taghavi (1996) generated several FE models with different element sizes. Once FE models were generated, the program BoneMat (Zannoni et al., 1999) was employed to assign to every element of the mesh the density value that was corresponding to its location. The relationship between Hounsfield units and density values were already calibrated using a phantom.

The obtained FE models were imported to ABAQUS/CAE. Only the coarsest model is discussed in this subsection. The two finer meshes and their application are discussed in a following subsection of the current section, namely Section 2.4. (convergence study). The bone geometry was already discretized using hexahedron elements (7934 elements, ABAQUS element type C3D8I, 8-node linear element, incompatible modes). The full geometry obtained from the VKH project was eventually replaced by a fraction of the femur (proximal femur, 3093 element) to reduce the required computational time (Fig. 1).

The elastic modulus at every location was calculated from the given density according to the following relationship:

\[ E = C\rho^\gamma. \]  

Schileo et al. (2007) compared the different values of the parameters C and \( \gamma \) that are reported in different studies to
determine which values can best describe the mechanical properties of the femoral bone. Based on the conclusions of their study, the parameter values reported in Morgan et al. (2003) \((C = 6950 \text{ and } \gamma = 1.49)\) were selected for relating the density to the elastic modulus of the femur \((174 \text{ different material definitions})\). Poisson’s ratio was fixed at 0.3 for all density values.

The hip contact force (HCP) was represented as a traction load distributed over a few elements. The traction load was specified in terms of force per area. Force per area is called traction intensity (traction pressure). The magnitude and the angle of application of the load in the frontal plane (about \(z\) axis, Fig. 1) were the parameters to be identified using the ANN. In order to have an estimation of muscle loads during the gait cycle, the muscle forces reported in a study by Bitsakos et al. (2005) were used. It should be noted that muscle forces are different from one person to another and change in the different parts of the gait cycle. Given the fact that the general trend of bone tissue adaptation models are found to be the same for different values of muscle forces (Bitsakos et al., 2005), we assumed that for the purpose of the current demonstrative study, the forces reported in Bitsakos et al. (2005) could be used for the cadaveric femur used here. The muscle forces were further assumed to remain unchanged during the different parts of the gait cycle. All the forces (Table 1, 10% gait cycle) were scaled to a bodyweight (BW) of 735 N, following the loading patterns proposed by Stolk et al. (2001) and Duda et al. (1998). Muscles with a wider insertion area were split into two parts.

At the distal end of the geometry, the nodes on the periphery of the diaphysis were constrained such that they could move in the \(x\) direction (Fig. 1). Three nodes were pinned on the side of the diaphysis to remove rigid body movements. Even though the constraints may influence the stress distribution in the areas close to the distal end of the geometry, according to the Saint Venant’s principle, they do not affect the load distribution of the femoral head.

---

**Table 1 – The muscle forces used in Fig. 1 (adopted for the case of 10% gait step from Bitsakos et al. (2005)). The values of the muscle force are presented in terms of the percentage of the Body Weight (BW). The orientation of the muscle forces are specified in terms of the angles that the force vectors make with the three axes of the reference coordinate system.**

<table>
<thead>
<tr>
<th>No</th>
<th>Muscle name</th>
<th>Magnitude (% BW)</th>
<th>Angles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(X)</td>
</tr>
<tr>
<td>1a</td>
<td>Gluteus maximus part 1</td>
<td>18.3</td>
<td>60</td>
</tr>
<tr>
<td>1b</td>
<td>Gluteus maximus part 2</td>
<td>14.5</td>
<td>58</td>
</tr>
<tr>
<td>2a</td>
<td>Gluteus medius part 1</td>
<td>14.0</td>
<td>59</td>
</tr>
<tr>
<td>2b</td>
<td>Gluteus medius part 2</td>
<td>23.2</td>
<td>44</td>
</tr>
<tr>
<td>3</td>
<td>Gluteus minimus</td>
<td>8.9</td>
<td>39</td>
</tr>
<tr>
<td>4a</td>
<td>Tensor fasciae latae 1</td>
<td>6.2</td>
<td>67</td>
</tr>
<tr>
<td>4b</td>
<td>Tensor fasciae latae 2</td>
<td>6.2</td>
<td>127</td>
</tr>
<tr>
<td>5</td>
<td>Psoas major</td>
<td>0.0</td>
<td>66</td>
</tr>
<tr>
<td>6</td>
<td>Iliacus</td>
<td>0.0</td>
<td>105</td>
</tr>
<tr>
<td>7</td>
<td>Piriformis</td>
<td>12.0</td>
<td>40</td>
</tr>
<tr>
<td>8</td>
<td>Quadratus</td>
<td>6.5</td>
<td>51</td>
</tr>
<tr>
<td>9</td>
<td>Abductor externus</td>
<td>8.0</td>
<td>54</td>
</tr>
<tr>
<td>10</td>
<td>Abductor internus</td>
<td>8.0</td>
<td>53</td>
</tr>
<tr>
<td>11</td>
<td>Gemellus</td>
<td>8.0</td>
<td>104</td>
</tr>
</tbody>
</table>
2.2. Tissue adaptation model

A strain-based tissue adaptation model (Weinans et al., 1992) is used in the current study. In this model, the change in the bone density is linked to the first principal strain, $\varepsilon_1$:

$$\Delta \rho = B(\varepsilon_1 / \rho - k) \Delta t. \quad (2)$$

The reference stimulus, $k$ is assumed to equal $0.0001 \text{ cm}^3/\text{g}$. The parameter $B$ regulates the speed of remodeling. The results of the bone tissue adaptation model were found to be largely independent from the choice of the parameter $B$. To speed up calculations, a $B$ value of $10^4 \text{ g}^2/\text{cm}^6 \text{s}$ was used in the current study.

The tissue adaptation problem is solved using an implicit FE solver (ABAQUS) in an iterative procedure. The simulation starts from a homogeneous density distribution, meaning that the density of all elements of the FE model is the same ($=0.8 \text{ g/cm}^3$). In every iteration, the FE model (Section 2.1) calculates the stress and strains resulting from the applied load using the current values of bone density. For every element of the FE model, the first principal strain as well as the current density is sent to a Fortran subroutine that uses Eq. (2) to calculate the change in the density of the element. The density and, thus, mechanical properties of the elements (Eq. (1)) are accordingly updated. The tissue adaptation model was solved for 30 iterations after which no significant changes in the density were observed.

2.3. Neural network load prediction

A schematic drawing of the method proposed for estimating the loads from density distribution is presented in Fig. 2(a). The whole procedure is composed of two main components, namely wavelet decomposition and ANN.

In the first component, a wavelet decomposition technique is used to reduce the number of inputs to ANN. In principle, the inputs of the ANN may be the density values of all the elements of the FE model (3093 inputs) without any modifications. This will, however, result in a large number of inputs, meaning that the computational cost will be high. In the current study, we proposed that multi-level wavelet decomposition could be used to reduce the number of inputs while preserving the most important information. In this approach, the vector containing the density values of the elements of the FE model is treated as a signal. A signal decomposition method is borrowed from the signal processing community and is applied for breaking down the density signal while preserving the important information required by the ANN for distinguishing between different density distributions that result from different loading parameters.

Here we only present a brief description of the wavelet decomposition technique. The reader is referred to Mallat (1989) for details of the wavelet decomposition method and the related algorithm. Wavelet transformations can be seen as generalizations of the Fourier transform. Instead of relying...
on sine and cosine functions as their only basis functions, wavelet transformations can be done using a variety of basis functions including the Haar function. A schematic drawing of the decomposition algorithm is presented in Fig. 2(b). Using wavelet transformation, one can extract most of the information out of the signal, while reducing its size (signal compression). In this approach, the signal, $S$, is decomposed into an important part ($A_1$, approximation signal) and a non-important part ($D_1$, details signal). The decomposition is performed using a low-pass and a high-pass filter. In a multi-level decomposition scheme, the approximation signal itself can be decomposed into $A_2$ and $D_2$ to reduce the size of the signal even further. Every time the signal is decomposed, the size of the signal is halved. Therefore, a signal that has gone through $n$-level decomposition is $2^n$ times smaller than its original size.

The density values of the training dataset were therefore first decomposed at the fourth level using a multilevel one-dimensional wavelet decomposition algorithm that used the Haar wavelet basis function. The number of inputs was thereby reduced from 3093 (the number of elements in the FE model) to 194 (the number of entries of fourth-level decomposed density signal).

The second component of the proposed method is an ANN. ANNs can be seen as nonlinear mappings from the space of inputs to the space of outputs. The reader is referred to the open literature (Hambli, 2011; Hambli et al., 2010; Hopfield, 1988; Wilamowski, 2009) and to our previous methodological paper (Zadpoor et al., submitted for publication) for detailed information on the structure and function of ANNs.

There are different types of ANNs that can be used for different computational tasks including feedforward, radial basis function (RBF), counterpropagation, and learning vector quantization (LVQ) networks (Wilamowski, 2009). It mathematically proven that feedforward ANNs with at least one hidden layer and sigmoid activation functions can approximate any continuous function regardless of the dimension of the input space with an integrated square error of the order $(1/n)$ (Barron, 1993). Here, $n$ is the number of the hidden neurons. A schematic drawing of feedforward ANNs is presented in Fig. 2(c). In the current study, we used feedforward ANNs with one hidden layer, one input layer, and one output layer. The hidden neurons were activated using a tangent-sigmoid activation function. The number of input parameters is the same as the size of the vector of decomposed density distribution. The outputs of the ANN are the magnitude and the angle of application of the hip contact force (2 outputs). The ANN has to be taught (trained) how to do the mapping from the space of density distribution to the space of the loading parameters. Solving the forward tissue adaptation problem for a large number of loading parameters generated the dataset that is required for training the ANN. The load magnitude varied between 0.11 and 2.31 BW with 0.044 BW distance between two consecutive load magnitudes. The angle of load application varied between $195^\circ$ and $225^\circ$ with $1^\circ$ distance between two consecutive angles. The loading parameters and (wavelet-decomposed) density distributions are known for the entries of the training dataset (1581 combinations of loading parameters). During the training process, the training algorithm adjusts the internal parameters of the ANN such that when the density distributions present in the training dataset are introduced to the ANN as input, the outputs of the ANN are as close to the corresponding loading parameters as possible. The ANN was trained using a Levenberg-Marquardt backpropagation algorithm. During the training phase, the training error was defined as the Mean Squared Error between the target values of the loading parameters and ANN-identified values of the loading parameters.

The training of the ANNs was in most cases carried out using 80% of the training data points. The remaining 20% of the training data points were divided into two subsets that were used for the validation (15%) and testing (5%) of the trained ANN. The validation dataset is not used in the training of the ANN, but is used to evaluate the improvement of the generalization capability of the ANN while it is being trained. The test dataset is not used at all in the training process and can therefore be used for assessment of the capability of the ANN in prediction of the loading parameters from the density distributions that are never seen by the ANN before. The results of a preliminary study showed that 10 hidden neurons are enough for accurate estimation of the loading parameters. If the values of some of the output parameters of an ANN (e.g. load and angle) are different by several orders of magnitude, the accuracy of the network may be compromised. Scaling is a way for avoiding such problems. In scaling, the values of one or more outputs of the ANN are uniformly scaled such that all the values of the output parameters are within a certain pre-defined range. In this study, the angle of load application was divided by 500 to bring the range of both loading parameters (traction intensity and angle) within the [0, 1] range. On a single core of an Intel Core i7 CPU, every run of the forward model (i.e. the FE bone adaptation model) takes about 65 s. Using three of the available cores of the same CPU, the total running time was less than 10 h. The training process of the ANN takes less than a minute.

2.4. Convergence study

The initial study was carried out using a relatively coarse mesh (3093 elements, see Section 2.1). The use of that relatively coarse mesh facilitated our exploratory study during which we examined various combinations of loading parameters and different ANN designs. Once the details of the methodological approach were finalized based on the exploratory study, we had to show that it is a robust methodology and that the loading parameters estimated with that proposed methodology converge and are not mesh-sensitive. In order to confirm the convergence of the obtained results, we included two additional FE models with finer meshes and called them the medium (7540 elements) and fine (41 624 elements) models. In that context, the FE model discussed in the previous sections is called the coarse model (3093 elements). A similar methodology as described in Sections 2.1–2.3 was applied to the medium and fine FE models. The major difference was the larger number of elements and, thus, larger number of inputs for the ANN. The forward adaptive simulations were run in parallel using 4 cores of the above-mentioned CPUs. The computation time of the medium and fine FE models were, respectively, 77 and 555 s. For comparison, the computational time of the coarse model was 42 s when it was run on 4 cores of the CPU.
3. Results

The results are presented in three subsections. The first two subsections (Sections 3.1 and 3.2) concern the coarse model, whereas the last subsection (Section 3.3) presents the results of the convergence study. As for the coarse model, two different approaches are used for the training of the ANN. In the first approach (Section 3.1), the training data generated by the forward adaptive model are used for the training of the ANN without any modifications. In the second approach (Section 3.2), the robustness of the ANN is improved by introducing noise to the training data.

3.1. Noise-free training of the ANN

The ANN trained using the generated training dataset was found to be able to accurately predict the load magnitude and angle of application for all the three subsets of the training datasets, namely training, validation, and test subsets (Fig. 3(a)–(c)). In Fig. 3, the horizontal axis shows the actual loading parameters (target values, \(T\)) while the vertical axis shows the loading parameters that were identified by the ANN (output of the ANN, \(Y\)). In the ideal situation, all the outputs of the ANN are the same as their corresponding target values, meaning that all the data points lie on the \(Y = T\) line. The Pearson correlation coefficient, \(R\), between the vector containing the actual loading parameters and the vector containing identified loading parameters was calculated (\(R\)-values are presented on top of the subfigures of Fig. 3). The Pearson correlation coefficients were greater than 0.999 for all the subsets of the training dataset, indicating the excellent capability of the ANN in prediction of the load from simulated density distributions. The absolute identification error was calculated as the absolute value of the difference between the identified loading parameters and actual loading parameters. The distribution of absolute identification error is not far from a normal distribution (Fig. 3(d)) as also implied by a small mean identification error (Table 2). The standard deviation of identification error and the mean absolute identification error (Table 2) were also small, further supporting the conclusion that the ANN can accurately identify the loading parameters from simulated density distributions. The next step is to use the trained ANN for identification of the loading parameters from the CT-measured density distribution (Fig. 4, see Section 2.1). The CT-measured density distribution was first decomposed using the above-mentioned wavelet decomposition technique and was then introduced to the ANN. The loading parameters were identified for two different training trials of the ANN. The identified loading parameters were consecutively used for simulation of the forward tissue adaptation model (Fig. 5). Comparing Figs. 4 and 5, one can see that the identified loading parameters and, as a result, the predicted density distributions are not accurate at all. In order to quantify the difference between the CT-measured density distribution and the density distribution predicted by the forward adaptive model, a number of error quantifiers were calculated (Table 3, rows 1 and 2). The error quantifiers were calculated based on the absolute difference between the CT-measured, \(\rho_{CT}\), and predicted, \(\rho_{FE}\), density values for every element, \(i\), of the FE model:

\[
\text{Error}_i = \left| \frac{\rho_{CT} - \rho_{FE}}{\rho_{CT}} \right| \times 100.
\]
The mean and standard deviation of the error values were calculated for two different sets of elements: the first set of elements contained all the elements of the FE model (volume error), while the second set of elements only consisted of the elements that belonged to the middle cross section (Fig. 4) of the FE model (cross section error). It was observed that the error values, as expected, are quite large (up to \( \approx 35\% \)).

It was therefore concluded that even though the ANN works very accurately for simulated density distributions (Fig. 3(a)–(c)), it is not capable of identifying the loading parameters from an actually measured density distribution.

### 3.2. Training of the ANN with noisy training data

We hypothesized that this may be due to the deviations of the CT-measured density distribution from the density distribution predicted by the forward tissue adaptation model. Since the ANN is trained using only simulated density distributions that are very accurately conforming to the tissue adaptation algorithm, it is not able to handle deviations from the predictions of the tissue adaptation theory that exist in the actual measurement of the density distribution. In order to make the ANN more robust against such deviations, a second ANN was trained this time using a noisy training dataset. Noise was introduced to the wavelet-decomposed density distributions, \( \rho_{\text{Haar},4} \), of the training dataset by adding a Gaussian random term as follows:

\[
\hat{\rho}_{\text{Haar},4}(x) = \rho_{\text{Haar},4}(x) + G(\rho_{\text{Haar},4}(x)/\lambda)
\]

where \( G \) is a Gaussian distribution and \( \lambda \) stands for the signal to noise ratio. The training procedure used for the second ANN was the same as the one used for the first ANN. Different values of signal to noise ratio varying between 5 and 50 were used. For every signal to noise ratio, a separate ANN was trained. The CT-measured density distribution was then introduced to the trained ANN. The identified loading parameters were used to run the forward tissue adaptation model. The resulting density distribution was compared with the CT-measured density distribution. The signal to noise ratios between 5 and 10 generated the best results (least error), with an optimal signal to noise ratio value around 7. A signal to noise ratio of 7 was therefore detected as the best signal to noise ratio and was chosen for further analysis. One can see that the ANN trained using noisy training dataset (Fig. 6(a)–(c)) can identify the loading parameters from simulated density distributions with a good accuracy. The distribution of absolute identification error is somewhat skewed (Fig. 6(d)) when compared with normal distributions. Most absolute identification errors are, however, concentrated in two small intervals on both sides of zero (Fig. 6(d)). Comparing the statistical parameters reported for the two different ANNs (Table 2), one could see that the mean and standard deviation of identification errors as well as mean absolute identification error are larger for the ANN trained with a noisy training dataset. The forward simulation of the tissue adaptation model using the loading parameters identified using the second ANN (Fig. 7) shows that the identified loading parameters result in a density distribution that is much closer to the measured density distribution (compare Figs. 5 and 7 with Fig. 4). The error values that are calculated for the new ANN (Table 3, row 3) are also lower than the ones calculated for the previous ANN.

### 3.3. Results of the convergence study

Simulations carried out for the medium and fine FE models show that the loading parameters identified for both finer models are close to the loading parameters identified for the coarse model (Table 3, compare rows 4-5 with row 3). However, one can see that the load magnitude slightly decreases and the angle of load application slightly increases as the number of elements increases. The forward adaptive simulations that were run using the identified loading parameters of the finer meshes (Fig. 8) show that the predicted density distributions are also similar to the one...
Table 2 – Statistical parameters calculated for the identification errors of both ANNs. The samples used for error analysis were not used in the training process at all.

<table>
<thead>
<tr>
<th>Type of training</th>
<th>Parameter</th>
<th>$\mu_{err}$</th>
<th>$\sigma_{err}$</th>
<th>$\mu_{abs}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise-free training</td>
<td>Traction intensity (BW/mm²)</td>
<td>1.077e-04</td>
<td>0.0023</td>
<td>0.0019</td>
</tr>
<tr>
<td></td>
<td>Angle/500 (deg)</td>
<td>-0.0015</td>
<td>0.0084</td>
<td>0.0047</td>
</tr>
<tr>
<td>Noise introduced</td>
<td>Traction intensity (BW/mm²)</td>
<td>0.0015</td>
<td>0.0032</td>
<td>0.0028</td>
</tr>
<tr>
<td></td>
<td>Angle/500 (deg)</td>
<td>-0.0093</td>
<td>0.0128</td>
<td>0.0109</td>
</tr>
</tbody>
</table>

Table symbols: $\mu_{err}$: mean identification error, $\sigma_{err}$: standard deviation of identification error, $\mu_{abs}$: mean absolute identification error.

Table 3 – The loading parameters identified by all the ANNs used in the current study are presented in this table. The mean and standard deviation of the absolute values of prediction errors are also presented. The error values are presented for two cases: the case where all the elements of the models are included in the error calculation (volume error) and the case where only the elements belonging to the middle cross section (shown in Figs. 4, 5, 7 and 8) are included in the error calculation.

<table>
<thead>
<tr>
<th>Name</th>
<th>Identified loading parameters</th>
<th>Volume error</th>
<th>Cross-section error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Magnitude (BW)</td>
<td>Angle (°)</td>
<td>µ</td>
</tr>
<tr>
<td>Coarse (without noise)</td>
<td>1.86</td>
<td>59.7</td>
<td>20.42</td>
</tr>
<tr>
<td>Coarse (with noise)</td>
<td>-0.89</td>
<td>132.4</td>
<td>27.09</td>
</tr>
<tr>
<td>Medium</td>
<td>1.67</td>
<td>211.2</td>
<td>16.17</td>
</tr>
<tr>
<td>Fine</td>
<td>1.52</td>
<td>212.0</td>
<td>16.80</td>
</tr>
</tbody>
</table>

Table symbols: $\mu$: mean of the errors values calculated for different elements of the FE models Eq. (3), $\sigma$: standard deviation of the errors values calculated for different elements of the FE models Eq. (3).

4. Discussion

Artificial neural networks were successfully used here to predict the loading parameters from a density distribution measured at the organ level. Based on the test results of the first ANN (Fig. 3 and Table 2), it can be concluded
that the ANN can predict the loading parameters using the wavelet-decomposed density distribution, meaning that the information required for distinguishing between the density distributions resulting from different loading parameters is preserved after wavelet decomposition. This is an important point, because wavelet decomposition drastically (>93%) reduces the number of the inputs of the ANN and results in reduction of the required memory and computational power.

Even though the first ANN could predict the loading parameters from simulated density distributions (Fig. 3), it failed to predict the correct loading parameters from the measured density distribution (Fig. 5). It was hypothesized that this difference is due to the fact that the ANN is trained using simulated density distributions that are very accurately conforming to the deployed tissue adaptation model. The density distribution measured using computed tomography obviously does not conform to the tissue adaptation algorithm as closely as the simulated density distributions are. This can be due to several reasons. First, the measurement of the density distribution using CT is not perfectly accurate (Hangartner and Gilsanz, 1996; Prevrhal et al., 1999; Treece et al., 2010). Second, the measured density distribution has not only resulted from mechanical loads. Biological factors also play an important role. An ANN that is merely trained using mechanical loading data cannot describe all the variations of the measured density distribution. Third, the tissue adaptation algorithm used here, and for that matter any other tissue adaptation algorithm, is simply a speculative idealization of the biological tissue adaptation process. It is not clear to what extent the deployed tissue adaptation algorithm can capture the actual biological process of bone tissue adaptation.

Fourth, the relationships that are used for relating bone density to the mechanical properties of bone are not perfectly accurate. Moreover, there are variations between different CT scanners and scanner settings that may influence how the gray values (Hounsfield units) are related to bone density and bone mechanical properties. Not all those variations are eliminated in the calibration process that is normally used to minimize them.

Finally, the number and magnitude of the loads that are used for simulating the forward tissue adaptation model are, to different degrees, different from actual physiological loads applied to the different parts of the femur. For example, the orientation and magnitude of all muscle loads change during the gait cycle, whereas the magnitude and orientation of all muscle loads were assumed to remain unchanged in the FE model used in the current study. Due to all the aforementioned reasons, it is expected that the density distribution measured by CT cannot be completely described using the forward tissue adaptation model. In order to improve the robustness of the ANN in dealing
with such deviations from the forward model, the ANN has to be exposed to some deviations from the forward tissue adaptation model in the training dataset (Zadpoor et al., submitted for publication). Testing the second ANN that was trained using a noisy training dataset (Figs. 6 and 7) showed that introducing noise to the training dataset increases the robustness of the ANN to a great extent.

As already mentioned, the best results were obtained for signal to noise ratios between 5 and 10, meaning that the deviations from the predictions of the forward tissue adaptation model are between 10% to 20%, with the best results obtained for about 15% deviation ($\lambda = 7$). The accuracy limit of CT in measurement of the density of cortical bone is reported to be around 10% (Prevrhal et al., 1999). The deviations from the prediction of the forward tissue adaptation models ($\lambda = 7$) are therefore quite close to the accuracy limit of CT densitometry. More advanced techniques such as peripheral quantitative computed tomography (pQCT) yield better accuracies (Augat et al., 1998). These techniques were, however, not used in the data collection protocol of the VAKHUM project. In order to separate the effects of the accuracy of CT measurement from other effects (discussed above), it would be interesting to study the effects of different CT techniques in the level of noise required for obtaining accurate prediction of the loading parameters.

The density distribution resulted from the loading parameters predicted by the second ANN (Fig. 7, Table 3) is much closer to the CT-measured density distribution as compared to the density distribution resulting from the loading parameters identified by the first ANN (Fig. 5, Table 3). The most important characteristics of the measured density distribution such as existence of high-density cortical bone around diaphysis and high density values in an area within the femoral head exist in both predicted (Fig. 7) and measured (Fig. 4) density distributions. There are, however, two major deviations from the measured density distribution in two particular areas of the simulated density distribution. First, relatively larger density values can be seen in the most medial part of the femoral head of the measured density distribution (Fig. 4) as compared to the predicted density distribution (Fig. 7). This is due to the fact that the femur contact force is represented by one single load in the current study. In reality, the load is distributed over a larger area of the femoral head. The second deviation occurs at the most distal part of the geometry. In this part of the geometry, the predicted density values are lower than the measured density values. This effect is most probably due to the boundary effects that are caused by the application of boundary conditions at this location.

The identified value of load per surface area was multiplied by the surface area of the applied load (110.9 mm$^2$) to obtain the magnitude of the femoral load in terms of BW. The total hip contact force predicted by the second ANN, i.e. 1.73 BW, is within the range of minimum and maximum hip contact forces reported in the literature (Bergmann et al., 2001; Heller et al., 2001).

The results of the convergence study showed that neither the estimated loading parameters nor predicted density distributions drastically change, even when the number of elements increases more than 13-fold from 3093 to 41624. The difference between the CT-measured and predicted density distributions are also not much different between the different FE models. However, one can see that the finer models overestimate the density values in certain area of the geometry (Fig. 8(c)–(d)) as compared to the coarsest model (Fig. 7). Those areas are the places where muscle or joint loads are applied. Application of loads affects the local density values less when the size of the elements is larger.

4.1. Limitations

The current study has some limitations. First, only one femur was used in the study. It would be more statistically powerful and interesting to apply the method presented here to a larger population of proximal femora.

Second, the number of loads used for representing the hip contact force was limited to one load (in the x-y plane). The density distribution predicted by the model is therefore only accurate in the mid frontal plane of the femur and the out-of-plane effects of the three-dimensional loads that are experienced by the femur cannot be fully captured by the model. In order to be able to capture the out-of-plane effects of the loads, one has to use several loads. Moreover, increasing the number of forces representing the hip contact force would enable it to better capture the density distribution of the other parts of the femoral head. If the number of loads is increased, the number of parameters that need to be identified by the ANN will increase. As a result, the computational cost and number of training samples that are needed for the training of the ANN will significantly increase. The other aspects of the load prediction process are not expected to change.

Third, generic muscle forces were used in the current study. For a more accurate representation of muscle forces, one can use a subject-specific musculoskeletal model or, even better, identify the muscle loads in a procedure similar to the procedure that was used for identification of the hip contact force. Predicting muscle forces using an ANN would, however, require a large number of training data points, as the number of parameters that need to be identified increases significantly.

Despite these limitations and simplifications, the current paper shows the feasibility and applicability of the proposed ANN load prediction method for prediction of the load from three-dimensional organ-level density distribution using a forward bone tissue adaptation algorithm and FE models.

4.2. Recommendations for future research

The work presented in this paper can be extended in many directions. In this subsection, we discuss two of those directions, namely estimating a large number of loading parameters and applying the proposed methodology to other femora.

The loading parameters influence the density distribution in a coupled way; meaning that the contributions of different loads to the resulting density distribution cannot be separated from each other. In order to capture the coupled behavior of the different loading parameters, one may have to use a full-factorial (fully crossed) design for generating the matrix of the loading parameters for which the forward adaptive model should be run. Assuming that the loading
parameter, \( i \), can take \( n_i \) different (discretized) values, a full-factorial design would require \( N = n_1 \cdot n_2 \cdot n_3 \cdot \ldots \cdot n_m \) different combinations of training data points for covering the full space of \( n \) loading parameters. One can therefore see that adding one additional loading parameter will increase the computational cost in a multiplicative way. If running a full-factorial design is not logistically possible, one may resort to fractional-factorial designs where a fraction of all possible combinations are included in the training data set. The decision as to what combinations of loading parameters should be included in the fractional factorial design can be made based on such methods as orthogonal or nearly-orthogonal arrays that are shown to be efficient in capturing the coupled effects of a large number of parameters while minimizing the computational cost (Wang and Wu, 1992; Wu et al., 2000).

Even though the trained ANNs can estimate the loading parameters of the individual femur studied in the paper, it is not possible to use those ANNs for estimating the loading parameters of other femora. A new ANN has to be trained for estimating the loading parameters of any other femur. It is therefore necessary to generate the training data separately for every other femur that needs to be analyzed. Since generation of the training dataset is the most computationally expensive part of the proposed method, it may not be feasible to train a new ANN for every femur that needs to be analyzed. However, the proposed method can be combined with Statistical Shape and Intensity Models (SSIM) (Barratt et al., 2008; Heimann and Meinzer, 2009) of the femur in order to make the methodology more feasible for estimating the loading parameters of an arbitrarily shaped femur.

In that combined ANN-SSIM approach, the shape and density distribution of any given femur are approximated as a weighted linear combination of a certain number of shape and density modes. Those shape and density modes are calculated from a statistical population of bones, and are assumed to be the same for all the other subjects whose femoral shape and density distributions are going to be approximated using the SSIM. The main idea is to find a small set of modes that best describe the shape and density variations of bones. The number of required modes is dependent on the desired accuracy of the approximation. For example, one study showed that the first 8 and 35 modes, respectively, describe +75% and +95% of the variation. Suppose \( p \) modes are considered adequate for approximation of femora. One only needs to specify \( p \) parameters, \( a_1, a_2, \ldots, a_p \), in order to describe (approximate) the shape and density distribution of any given femur. Those parameters determine the relative influence of different modes in construction of the shape and density distribution of that femur.

The coupling between ANN and SSIM can be done in the following way. First, a large number of virtual femora are generated by varying \( a_i \) values. For every virtual femur, the forward adaptive model is run for a large number of loading parameters. Once all the data is generated, a new ANN is trained using the generated data such that the ANN receives both (decomposed) density distribution and \( a_i \) values as input and returns the loading parameters as output. Generating the required training data and training the above-mentioned ANN is a huge computational job that can only be accomplished feasibly using massive (super-)computing resources. However, that is a one-time computational job. Once trained and tested, the ANN can be used for estimating the loading parameters of any femur whose \( a_i \) values and density distribution are determined within a fraction of a second. It is important to note that a 3D matching algorithm is needed for determining \( a_i \) values. The 3D matching algorithm has to position and align the femur according to the position/alignment convention that is used for constructing the SSIM.

### References


