# COMPARISON METHODS FOR IMPROVING GENERALIZATION WHEN PREDICTING OF MUSCLE FORCES USING A NEURAL NETWORK

J. Vejpustková*, M. Vilímek* and M. Sochor*

* CTU in Prague, Faculty of Mechanical Engineering, Department of Mechanics, Laboratory of Human Biomechanics, Technická 4, 16607 Prague 6, Czech Republic

Vejpustk@biomed.fsid.cvut.cz

**Abstract: Methods for directly measuring muscle forces are not so far available and calculating muscle forces is difficult, because many muscles act cooperatively. However, orthopeadists, biomechanical engineers and physical therapists need to take muscle forces into consideration because joint contact forces, as well as muscle forces, need to be estimated in order to understand the joint and bone loading. Therefore an artificial neural network (NN) with a learning algorithm was used in order to predict the muscle forces in elbow joint. This paper describes collecting and preprocessing the data; suggestion and training the NN. In conclusion generalization performance was improved, using various methods for preprocessing the data.**

## Introduction

In biomechanics, it is one of important issues to study the distribution of muscle forces between individual muscles in order to understand the joint and bone loading. There exist 4 general methods for estimating the muscle forces during human movements (heuristic method based on statics or inverse dynamics which are based on simple assumptions for load sharing, an inverse dynamical approach involving processing of experimental motion data, modelling and static optimization to solve the muscle redundancy problem, an EMG-to-force processing approach, direct dynamical approach involving model-driven simulations of the movement task). These methods are difficult to use, the calculations take a long time. Therefore an artificial neural network (NN) with a learning algorithm is used in order to predict the muscle forces in elbow joint. The elbow joint was selected because it provides a good visual demonstration, and the movement is uniplanar and uniarticular. Current NN can be trained to solve problems that are difficult for conventional computers or human beings. The big advantage of NN is obtaining results without knowledge of the algorithm procedure or without full and exact information.

The proposed object of the neural network [12] simulated the cooperation of 7 musculotendon actuators in the elbow joint, four flexors: *m.biceps brachii c.longum* and *c.breve*; *m.brachialis*; *m.brachioradialis*; and three extensors: *m.triceps brachii c.laterale*, *c.mediale* and *c.longum* for 4 movement conditions (combination of fast and slow motion and unloaded and with weight). In our study there was attempt to study all input parameters which influence resulting muscle forces. For input parameters were taken 14 muscle properties (physiological characteristics of muscles participating in the particular joint mechanism, together with data about the movement and electric activity of the muscles).

This paper is focused on improving generalization performance, using various methods for preprocessing the data.

## Materials and Methods

Using the Matlab Neural Network Toolbox, the backpropagation NN with a learning algorithm was programmed. Standard backpropagation NN is a gradient descent algorithm, in which the network weights are moved along the negative of the gradient of the performance function. The term backpropagation refers to the manner in which the gradient is computed for nonlinear multilayer networks. Properly trained NN tends to give reasonable answers when presented with inputs that the network has never seen.

*The architecture of Neural network object:* The architecture of NN was the feedforward multilayer network, in this case consisting of three layers (input layer and two hidden layers followed by an output layer). The network object with 30 neurons in the 1st hidden layer and with 24 neurons in the 2nd hidden layer was proposed. Between input layer and 1st hidden layer and between 1st and 2nd hidden layer there were used sigmoidal transfer function – tansig. Multilayer network uses the sigmoidal transfer functions, because they are differentiable functions. Between 2nd hidden layer and ouput layer was used linear transfer function – purelin. Linear transfer function was used so the neural outputs took on any value. A schematic representation of NN object is shown in Figure 1.

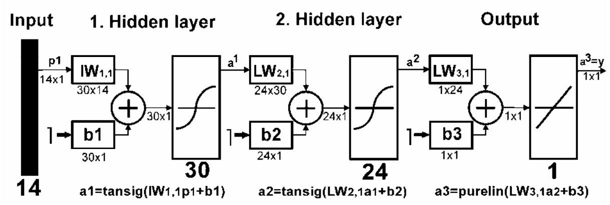


Figure 1: A schematic representation of three-layer feedforward NN with supervised learning algorithm.

*Input/target pairs:* In NN object, 14 input parameters were used for estimating the muscle forces. The used input parameters were: musculotendon length, $L_{MT}$, velocity of muscle shortening, $v$, pennation angle, $\alpha_0$, optimal muscle fibre length, $l_0$, physiological crossectional area, *PCSA*, tendon slack length, $L_{ST}$, maximal isometric muscle force, $F_0$, force-velocity factor, $F_v$, active force-muscle length factor, $Fl_a$, passive force-muscle length factor, $Fl_p$, muscle activity, $a(t)$, and three degrees of history of muscle activity, $a_{1H}(t+\Delta t)$, $a_{2H}(t+2\Delta t)$ and $a_{3H}(t+3\Delta t)$.

An arm movements were from full extension ($\varphi_E$=0°) to full flexion ($\varphi_F$=145°) [9] of the elbow joint for a fixed shoulder joint. The forearm was free to move in the sagittal plane of the elbow. The elbow flexion/extension movements were recorded using the 6-camera 60Hz VICON Motion Analysis system, two movement speeds (slow, 1.1rad/sec and fast, 2.8rad/sec) and two loading conditions (unloaded and with 4.2kg bar-bell) were studied. The electric activity of the observed muscles was recorded by surface (non invasive) electromyography (EMG). EMG is investigative method that is based on scan muscle activities. The processed EMG signal was done by filtering of frequences which are lower then 20Hz and higher then 500Hz, offsetting, rectifying (rendering the signal to have excursions of one polarity) and integrating the signal over a specified interval of time [1]. The processed and the normalized EMG signal was taken as the input of muscle activity, $a(t)$ and the history of muscle activity, $a_{1H}(t+\Delta t)$, $a_{2H}(t+2\Delta t)$, $a_{3H}(t+3\Delta t)$.

The muscles consist of an active force generating component and a parallel connective tissue component. The parallel connective tissue does not actively generate force but if it is stretched beyond its resting length produces a passivec force. As well as the musculotendon length, $L_{MT}$, having an effect on the maximum force it can generate, so does the velocity of muscle shortening, $v$. The musculotendon length, $L_{MT}$, (the length of the entire muscle-tendon unit origin to insertion) was estimated from anatomical positions of muscle attachments and recorded kinematic data, the velocity of muscle shortening, $v$, was calculated from recorded kinematic data (the slow movement and the fast movement unloaded, and loaded, respectively).

Some of the muscular parameters were taken from [11] (the optimal muscle fiber length, $l_0$, and the pennation angle, $\alpha_0$). Because *PCSA* crossing across all fibres of the muscle, was estimated the pennation angle, $\alpha_0$, which determines organization of fibres in a muscle. Tendon slack length, $L_{ST}$, was theoretically calculated by method published in [2] and [13]. Maximal isometric muscle force, $F_0$, was calculated as physiological crossectional area, *PCSA*, multiplied by specific muscle tension, σ, where specific muscle tension is $\sigma$=31.8N.cm$^{-2}$ [6]. The force-muscle length factor was taken into account in terms of [3], and the curves of passive, $Fl_p$, and active, $Fl_a$, properties, scaled to provide a destription for specific muscle are fit by parabolic and exponential functions Force-velocity factor, $F_v$, was calculated from hill equation [4] (for concentric contraction) and modified hill equation [7] (for eccentric contraction).

The output parameter (OP) must be known for training as well. As OP, the muscle force, applying the *Virtual Muscle system*, see [6], was used in order to relate this to the real muscle force. The *Virtual Muscle system* is a modeling package designed to run on Matlab (The Mathworks Inc., Natick, MA) and also requires Simuling module for Matlab. Inputs are muscle morphometrical data, time depending on length of musculotendon and muscle activity.

*Preprocessing the training data:* NN training was made more efficient if certain preprocessing steps were performed on the network representative set of input/target pairs. Post-training analyses were also carried out. The approach for scaling the network inputs and targets was to normalize the mean and standard deviation of the training set so that they had zero mean and unity standard deviation. Consequently, the dimension of the input vectors was reduced by principle component analysis (PCA) [5]. The input vectors were uncorrelated with each other and the components with the largest variation came first, which eliminated those components that contributed the least to the variation in the data set.

Two features of the Neural Network Toolbox were designed to improve network generalization - automated regularization (trainbr) [8] and the framework of early stopping.

In automated regularization, the weights and biases of the network were assumed to be random variables with specified distributions. The regularization parameters were related to the unknown variances associated with these distributions. When was used Bayesian regularization, it was important to train NN object until it has reached convergence.

In the second technique, early stopping, the preprocessing data was divided into three subsets. The first subset was the training set. The second subset was the validation set, which was used for computing the gradient and updating the network weights and biases. The third subset was used for simulation the network response. When the validation error increased, the training was stopped. The validation set should be representative of all points in the trainig set.

*Training NN object:* In the course of learning the NN object, the main goal was to find the solution with the smallest error and the fastest convergence. Minimization of learning error was performed by modifying the network topology, by changing the number of neurons in the hidden layers and by changing the learning rate. The NN object was also sensitive to the number of neurons in their hidden layers. Too few neurons led to underfitting. Too many neurons led to overfitting. If the learning rate of the network was set up too high, the correct solution was overskipped. If the network learning rate was too low, the correct solution very often ends in the local minimum, or the algorithm converges very slowly.

The training data was created from 7 musculotendon actuators about elbow joint for 4 movement conditions (combination of fast and slow motion and unloaded and with weight). 98 training data was considered for each movement. For each muscle 4 files were created, that means 98 training data x 4 movement = 392 training data. As a whole it was 7 muscles x 4 files = 28 files which representative all investigated movements. Finally, for the elbow problem was available 392 training data x 7 muscles = 2744 trianing data (all of the input sets from 7 elbow actuators).

**Results**

It was useful to investigate the network response in more detail, performed a regression analysis between the network response and the corresponding targets. The results of these methods (Table 1) show that the framework of early stopping with data preprocessed by principal component analysis (PCA) (C=0.97) provided the best results. The form of Bayesian regularization did not generalize well for a problem like prediction of muscle forces, see Figure 2. The results for trainig were excellent (C=0.999), but after generalization performance the results were not sufficient (C=0.917). Another disadvantage of the Bayesian regularization method was that it took longer to converge than early stopping.

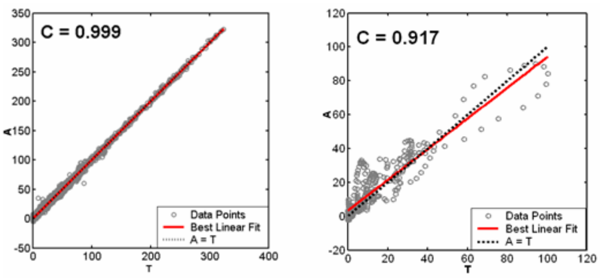


Figure 2: A regression analysis for Bayesian regularization between the network response and the corresponding targets. On the left side (learned data), C is very close to 1, which indicates a good fit. On the right side (generalized data), C is lower. The perfect fit (output equal to targets) is indicated by the dashed line.

Early stopping with data preprocessing performed better (C=0.97) than the early stopping without data preprocessing (C=0.89), see Figure 3. The response was less than extremely smooth, as when using regularization, but the error of the network response was smaller.

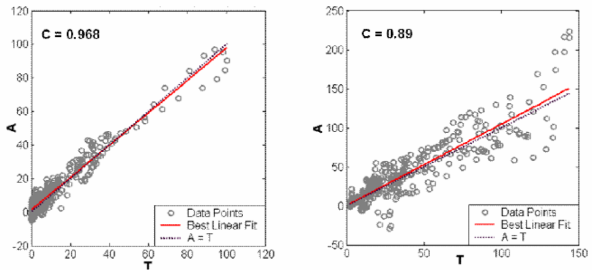


Figure 3: A regression analysis for early stopping between the network response and the corresponding targets. On the left side is processed data. On the right side is raw data. The perfect fit (output equal to targets) is indicated by the dashed line.

Table 1: In order to compare the methods for improving generalization when predicting the muscle forces using neural network, we used mean square error, MSE; mean absolute error, MAE; and correlation coefficient, C.

| **Error** | **Early stopping** | | **Bayesian regularization** |
|---|---|---|---|
| | **Processed data** | **Raw data** | |
| MSE[N] | 19.63 | 446.31 | 52.79 |
| MAE[-] | 3.46 | 13.45 | 4.55 |
| C[-] | 0.97 | 0.89 | 0.92 |

**Discussion**

This paper was focused on improving generalization performance, using various methods for preprocessing the data. For solution was used the neural network-type muscle object (NN object) which is a black-box without any attempt to follow the internal knowledge of the muscle komplex [10]. NN object depends on training sets and on current problems which are studied. NN it is a good instrument for tasks without knowledge of algorithm procedure or without full and exact information, for solving complicated mathematical description, ability of teaching and generalization. But on the other hand there are some disadvantages: difficult option of optimal network topology, complication of networks, long time for training and difficult estimation if generalization is correct, less suitable as universal instrument for exact calculations.

In every event using the artificial neural network to the muscle forces estimation is one of the possible ways and in the future it can be very strong technique for solving this type of problems.


ISSN: 1727-1983 © 2005 IFMBE

## Conclusions

Generally, this NN object could be used for predicting the muscle force for all muscles, not only for the elbow joint, but this of course depends on the training data and on preprocessing a representative set of input/target pairs. For good generalization and good results, it is best to use the framework of early stopping with data preprocessing with the help of principal component analysis (PCA). Early stopping performs well on pattern recognition problems such as prediction of muscle forces.

## Acknowledgements

The research has been supported by the Ministry of Education project: Transdisciplinary research in Biomedical Engineering II., No. MSM 6840770012.

## References

[1] DE LUCA C.L. (1997): 'The use of surface electromyography in biomechanics', *Journal of Applied Biomechanics*, 13, pp. 135–163
[2] GARNER B.A., PANDY M.G. (2003): 'Estimation of musculotendon properties in the human upper limb', *Annals of Biomedical Engineering*, 31, pp. 207–220
[3] GORDON A.M., HUXLEY A.F. (1966): 'The variation in isometric tension with sarcomere length in vertebrate muscle fibres', *Journal of Physiology*, 184, pp. 170-192
[4] HILL A.V. (1970): 'First and last experiments in muscle mechanics', Cambridge University Press, Cambridge
[5] HOTELLING H. (1933): 'Analysis of a Complex of Statistical Variables with Principal Components', *Journal of Educational Psychology*.
[6] CHENG E.J., BROWN I.E. and LOEB G.E. (2000): 'Virtual muscle: a computational approach to understanding the effects of muscle properties on motor control', *Journal of Neuroscience Methods*, 101, pp. 117–130
[7] KRYLOW A.M., SANDERCOCK T.G. (1997): 'Dynamic force responses of Muscle involving eccentric contraction', *Journal of Biomechanics*, 30, pp. 27–33
[8] MAC KAY D.J.C. (1992): 'Bayesian interpolation', *Neural Computation*, **4,** pp. 415–447
[9] RADFORD M., CARR A. (2002): 'Total elbow replacement', *Current Orthopaedics* 16, pp. 325-330.
[10] ROSEN J., FUCHS M.B., ARCAN M. (1999): 'Performances of Hill-type and neural network muscle models-toward a myosignal-based exoskeleton'. *Computers and Biomedial Research*, 32, pp. 415-439
[11] VEGER H.E.J., YU B., AN K.N. and ROZENDAL R.H. (1997), 'Parameters for modeling the arm', *Journal of Biomechanics*, 30, pp. 647–652
[12] VEJPUSTKOVÁ J., VILÍMEK M. (2004): 'A muscle forces prediction using learning artificial neural network', Biomechanics of man 2004, Plzeň: University of West Bohemia, 2004, pp. 395-401.
[13] VILÍMEK M. (2004): 'A Theoretical Conditions for Calculation of the Tendon Slack Length', In Summer Workshop of Applied Mechanics, 3, pp. 336–341


ISSN: 1727-1983 © 2005 IFMBE