A neural networks approach to determine factors associated with self-reported discomfort in picking tasks

Olfa Haj Mahmoud, Charles Pontonnier, Georges Dumont, Stéphane Poli, Franck Multon

A neural networks approach to determine factors associated with self-reported discomfort in picking tasks

Olfa Haj Mahmoud[1,2], Charles Pontonnier[2], Georges Dumont[2], Stéphane Poli[1] and Franck Multon [3]

Précis: An approach based on neural network for estimating reported discomfort of subjects on workstations using biomechanical, anthropometric and environmental variables. A new approach of input reduction was proposed while maintaining the accuracy of the prediction model.

Abstract

Objective: A neural networks approach has been proposed to handle various inputs such as postural, anthropometric and environmental variables in order to estimate self-reported discomfort in picking tasks. An input reduction method has been proposed, reducing the input variables to the minimum data required to estimate self-reported discomfort with similar accuracy as the neural network fed with all variables.

Background: Previous works have attempted to explore the relationship between several factors and self-reported discomfort using observational methods.The results showed that this relationship was not a simple linear relationship. Another study used neural networks to model the function returning reported

[1] FAURECIA Automotive Seating, Etampes, France
[2] Univ Rennes, CNRS, Inria, IRISA-UMR6974, M2S, 35042 Rennes,
[3] Inria, Univ Rennes, CNRS, IRISA-UMR6974, M2S, 35042 Rennes, France
Email: olfa.haj-mahmoud@irisa.fr

discomfort according to static posture, age and anthropometrics variables. The results demonstrated the model's ability to predict reported discomfort. But all the available variables were used to design the neural network.

Method: Eleven subjects carried-out picking tasks with various masses (0, 1, 3 kg) and imposed duration (5, 10 or 15 seconds). Continuous REBA score, anthropometric and environmental data were computed, and subjects' discomfort were collected. The data set of this work consisted in the computed continuous REBA score, anthropometric, environmental data and collected subjects' discomfort.

Results: The results showed that the correlation between the estimated and experimental tested data was equal to 0.775 when using all the 14 available variables. After data reduction, only 6 variables were left, with a very close performance when predicting discomfort.

Conclusion: A neural network approach has been proposed to estimate self-reported discomfort according to a minimum set of postural, anthropometric and environmental variables in picking tasks.

Application: This method has the potential to support ergonomists in workstation designing processes, by adding discomfort prediction to virtual manikins' behaviors in simulation tools.

Keywords:REBA, Physical risk factors, CR10, Kinematics, WMSD.

Introduction

Perceived discomfort at work has been associated with work musculoskeletal disorders (WMSDs) (Das and Gangopadhyay 2011), a very common occupational problem affecting parts of the human body (i.e. muscles, joints, tendons, nerves, etc.) (Ando et al. 2000). Several risk factors could be involved in the occurrence of these WMSDs, such as work organization, psychosocial, biomechanical and anthropometric factors of work (Da Costa and Vieira 2010).

Several complementary methods have been introduced to assess these risk factors. For example, biomechanical risk factors are generally evaluated using postural scores (such as the Rapid Upper Limb Assessment (RULA) (McAtamney and Corlett 1993) or the Rapid Entire Body Assessment (REBA) (Hignett and McAtamney 2000)), or task-related methods (such as NIOSH method for handling

tasks (Walters et al. 1994)). These methods provide quantified values, where increasing score supposedly reflects increasing discomfort. Perceived discomfort is also generally reported through interviews, and quantified scales, such as the Borg Rated Perceived Exertion (RPE) and Category Ratio 10 (CR10) scales (Borg 1998).

To prevent WMSDs, the complex combination between those factors and its relation to the perceived discomfort (Ando et al. 2000; Liao and Drury 2000) has to be understood. For industrial tasks requiring moderate to high postural stress, such as picking tasks, this relationship is difficult to characterize (National Research Council and Institute of Medicine 2001; Bhatnager et al. 1985). Previous works demonstrated a relationship between physical exertion at work and discomfort (Bhatnager et al. 1985; Kee and Karwowski 2001a; Liao and Drury 2000), but also with other biomechanical factors, such as awkward postures and repetition. This relationship has mainly been observed when manipulating heavy loads (Li et al. 2009), whereas this relation was less clear for light loads.

Chihara et al. (Chihara et al. 2014) proposed to investigate the relationship between observational methods (RULA, REBA) and reported discomfort for manual handling tasks using 0 to 10 kg masses. The results showed that the correlation between the perceived discomfort and observational methods was not significant, with a maximum correlation coefficient of 0.593. In a previous paper (Haj Mahmoud et al. 2020), we studied picking tasks for low masses (1 kg) using Principal Component analysis (PCA).The results showed no significant linear relationship between postural scores and reported discomfort for such tasks. On the opposite,Li et al. (Li et al. 2009) showed a significant linear relationship between RPE (Borg 1998) and physiological parameters (heart rate and oxygen uptake) for picking tasks, but with heavy loads (23 kg).

On the basis of those results, we can assume that the relationship between postural variables and reported discomfort is not limited to a linear correlation. It also depends on other factors, such as the load which is manipulated or the anthropometrics of the operator. Hence, the actual function to predict reported discomfort should look like $CR10 = f(REBA, load, etc.)$ instead of a linear function $CR10 = \alpha REBA + \beta$.

This function f has been studied in previous works using linear regression equations. It has been tested to explain the relationship between perceived comfort ratings and different levels of joint angles in sitting and standing postures (Kee and Karwowski 2001b). This study showed that the static posture may be one of the most significant factors affecting perceived joint discomfort, but they did not explicitly design a function to model the relationship between all the variables.

Other works considered using fuzzy logic to model the link between perceived discomfort, anthropometric parameters and task features. It has been applied to model driver's perceived discomfort when handling interior controls, using anthropometric parameters and control positions (Hanson et al. 2003). The results of this study showed that the performance of the achieved model seemed to be reasonable and relevant in this specific case.

All the previous approaches considered linear relationships between the variables. For static postures, Chryssolouris et al. (Chryssolouris 2004) demonstrated that a non linear approach such as a neural network was able to correctly estimate the reported discomfort according to postural parameters, age and anthropometric parameters (stature and body mass). However, the authors did not try investigate the impact of each variable on the accuracy of the prediction model.

To summarize, when trying to estimate the function f using linear regression between postural variables and reported discomfort, the results obtained in previous works were not sufficient. This confirmed that there exist complex interactions between variables when attempting to estimate reported discomfort. Neural networks have shown promising results but the previous works used only static postures, and did not investigate the impact of the input parameters to predict reported discomfort. Moreover, previous works focused on few joints posture analysis only, and did not considered other task-related parameters, such as grasping height, load, asymmetry angles between the orientation of the subject and the object to be manipulated, etc.

Machine learning is a promising mean to explore these complex interactions between parameters. Recent deep learning methods may not be usable in that case, as it would require a very large amount of data for training, (Zhang et al. 2018), which is difficult to get experimentally. The heterogeneous nature

of the input data prevents the use of convolution networks as well. Auto-encoders may be used to encode compact non-linear interactions between the input (Georgiou et al. 2019), but the resulting latent space would be impossible to analyze for future ergonomic use. Classifiers, such as Support Vector Machine or Decision trees would have difficulties to handle the natural variability of subjective reported discomfort for the same given situation. In other domains, several recent works modeled the relationship between physical variables and subjective feeling of users using shallow neural networks. These classical neural networks have recently been used to estimate the acoustic behavior of induction motors (Jiménez-Romero et al. 2020), sound quality evaluation of the interior noise of vehicles (Steinbach and Altinsoy 2019) and noise evaluation inside vehicles (Ma et al. 2017).

In the current study, we proposed to explore the performance of such a shallow neural network to model the complex relationship between postural, anthropometric and environmental parameters for picking and placing tasks (not limited to static postures). We also wished to analyze the sensitivity of each parameter in this prediction. To this end, we used an input reduction method to identify the minimal set of variables able to correctly estimate the reported discomfort. The resulting estimator could be used as a decision support system for ergonomists and engineers when designing/adapting work conditions. It could be applied to a real situation or a digital mockup with virtual manikins, at the early stage of the design process and to predict reported discomfort of the future users.

Material and methods

We carried-out an experimentation in which subjects had to pick and place objects with various masses, locations and timing. For each task, we collected self-reported discomfort, as well as anthropometric data, environmental conditions, and motion of each subject. We designed an initial neural network that best captures the relationship between these variables and the reported discomfort. Less significant task-related variables were progressively discarded until it was not possible to design a neural network as accurate as the first one, with all the collected data.

Participants and protocol

Eleven healthy subjects (seven men and four women, age 26.7 ± 4.1, stature 171.63 ± 13.9cm, body mass 70.72 ± 9.07kg), participated in this study. Each subject signed an informed consent form prior to the experiment.

The experimental task consisted in picking an object (barbell) weighing 0 kg, 1 kg or 3 kg at a starting position, and placing/maintaining it to a final position with different duration. Preliminary studies showed that beyond 3 kg, sedentary subjects may fail to maintain such a mass for 15s in front of them without the appearance of fatigue. We used the same holding mass for all the subjects, independently of their anthropometric parameters, as we did not want to introduce heuristics on the interactions between load mass and anthropometric parameters.

Starting position was set at 75 cm height left (resp. right) for left-handed (resp. right-handed) subjects. A height of 75 cm was chosen as it corresponded to a supposed comfortable height, according to the French Norm NF X35-109 (AFNOR 2015). Five final positions were proposed. Three final positions were placed in front of the subject at three different heights: 0 cm (ground), 75 cm and 175 cm. The two others final positions were placed on the side of the subject (same side as the starting position) at two different heights: 0 cm (ground), and 175 cm height. For each final position, the subject was asked to hold his posture (object in the air, without contact with a support) during 5, 10 or 15 seconds. A screen placed in front of the subject displayed the mass of the object to pick, the final position, and the holding time of the trial to be performed. Once the subject reached the final position, an operator controlled the holding time and informed orally the subject to return to the starting position when the time was up. Figure 1 depicts the experimental setup.

Each possible combination of trial (final position, mass, maintaining time) was repeated three times in a randomized order. Finally, the subject had to perform 5 (final positions) x 3 (masses) x 3 (maintaining time) x 3 (repetitions) = 135 total randomized trials.

Prior to the experiment, subjects were familiarized with various situations, including the supposed worst ones (lowest/highest final positions, with longest maintaining times). During this familiarization
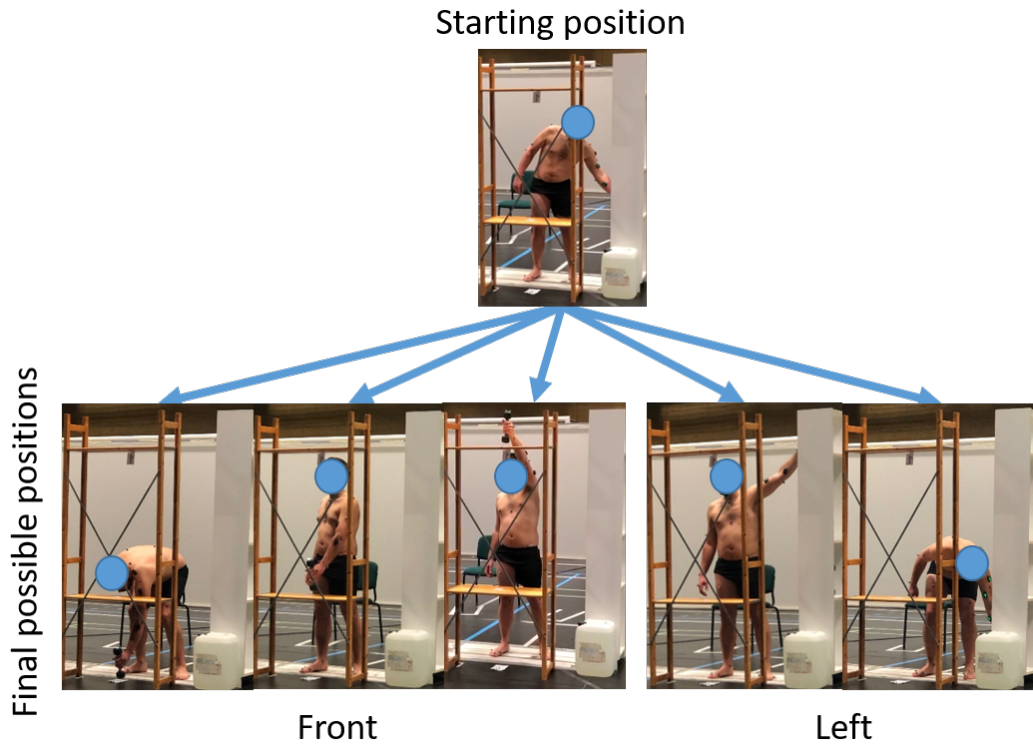
Figure 1. Experimental set-up with an object to pick at a unique starting position (top) and place at 5 possible final positions (bottom). The subject is left-handed in this example, so that the chest for picking and placing the object was placed on his left.

step, subjects were asked to assess their discomfort feelings (using 0-10 Borg's CR10 scale) for a reference task: placing the 0 kg object in front of the subject at 75 cm height.

After placing back the object on the shelf at the starting position, the subject had to rate discomfort using a 0-10 CR10 scale, as for the familiarization step. A short break was provided between trials to prevent from fatigue issues. The subject could also ask for a longer break at any moment. In any cases, we imposed a long break (5 minutes) 3 times along the protocol to prevent from fatigue or a loss of concentration.

Data collection

An optoelectronic motion capture system Qualisys (23 12-Mpixels cameras, 200Hz sampling frequency) was used to track 43 reflective markers placed on standardized anatomical landmarks, following the ISB (International Society of Biomechanics) recommendations (Wu and Cavanagh 1995). Based on the method proposed by (Plantard et al. 2017), we calculated the REBA score according to the ISB-compliant joint angles. The joint angles were used to compute the REBA postural score continuously along time for each trial. REBA method divides the body into body parts that are individually scored according to the joint angles: wrist, forearm, elbow, shoulder, neck, trunk, legs and knees. Then, the local REBA scores for each joint are combined to obtain a whole-body score. This whole body score is adjusted depending on the mass of the carried load, quality of grasp and frequency.

The motion of the subject was segmented to isolated picking and placing sub-tasks. Picking was starting when the object left the starting position, and was ending when it left the final position to go back to the initial place. Placing was starting at this time and was ending when the object was placed back to the initial position. Based on the continuous REBA score, we computed several task-related postural variables, as reported in (Haj Mahmoud et al. 2020): the mean, integral and Root Mean Square (RMS) values of the continuous REBA score, and the time spent at each REBA score level (expressed as time spent at REBA score 10 for *Score 10* variable, for example).

$$REBA_{Mean} = \frac{1}{N} \sum_{i=1}^{N} REBA(i) \tag{1}$$

$$REBA_{Integral} = \int_{i=1}^{N} REBA(i) d(i) \tag{2}$$

$$REBA_{RMS} = \sqrt{\frac{1}{N} (\sum_{i=1}^{N} REBA(i)^2)} \tag{3}$$

$$Score_k = Time(REBA = k) \tag{4}$$

With N the total number of frames, $REBA(i)$ is the value of REBA score at frame i.

We also collected the subject's body mass and stature. We finally collected environmental variables, as suggested in NIOSH-lifting equation (Walters et al. 1994): the height of the final position, the mass of the object to handle and the asymmetry angle (AA) defined as the trunk axial rotation.

Neural Network Method description

We used a shallow neural network to estimate the function estimating the reported discomfort based on the postural, anthropometric and environmental variables. Once a good estimation was obtained with this neural network, we applied an input reduction method to determine the minimum set of variables which enabled us to correctly estimate the reported discomfort.

Design of the neural network An artificial feedforward neural network (ANN) (Ruck et al. 1990), named Multi-Layer Perceptron, was used for its ability to capture nonlinear relationships between input x and output y variables. This kind of network consists of layers of interconnected neurons including activation functions, and all the connections have some mathematical expressions associated to weights. Similarly to previous works (Jiménez-Romero et al. 2020; Steinbach and Altinsoy 2019; Ma et al. 2017), we selected this classical type of neural network because of the heterogeneous nature of the input data, and the low number of examples that could be used for training. In this study, x was the set of variables collected during each trial (postural, anthropometric and environmental variables, as described in Table 1). y stands for the estimated reported discomfort.

We discarded variables with too few occurrences (almost no data available), such as the time spent at REBA score 1, which was not viewed during the experiment. y was the CR10 score associated with the reported discomfort. Thus, the goal of the neural network was to approximate the nonlinear function $y = f(x)$. As we collected 14 variables (size of x is 14), the first layer of the neural network was composed of 14 neurons. The last layer of the neural network was a unique neuron, as it computed the final estimation of y for each trial. The number of intermediate layers, and their amount of neurons,

Table 1. The various variables for posture, organizational and anthropometric factors used to estimate discomfort

Posture	Anthropometric	Environmental
$REBA_{Mean}$	Mass of body	Position of final height (5 levels)
$REBA_{Integral}$	Subject stature	Angle of asymmetry
$REBA_{RMS}$		Mass of object (0, 1, 3 kg)
Score 4		
Score 5		
Score 6		
Score 8		
Score 10		
Score 11		

was manually tuned to find the neural network with the best performance for all the collected variables, as proposed in (Chryssolouris 2004).

To implement this ANN, we used the Rapidminer Software (Markus Hofman and Ralf Klinkenberg 2014). Figure 2 depicts the neural network used to model the relationship between the 14 input variables and the CR10 estimation output.

As presented in the following section, the neural network was trained with a part of the collected data, using classical back-propagation algorithm, and tested on the remaining part of the data.

In our study, 75% of the data set was used to train the model, while 25% was used to test it, as reported in previous works (Okuyucu et al. 2007; Steinbach and Altinsoy 2019). The training and testing sets were selected randomly among the whole data set containing 1482 trials.

Validation process We tested the robustness of the proposed model and its performance using three validation methods. The first one was k-fold cross validation method that enabled us to test various distribution between training and testing data sets. The second one was bias and variance analysis to estimate the ability of the network to reproduce human performance (error and generalization capabilities). The third was a leave-one-out cross validation by eliminating all the data of one subject.
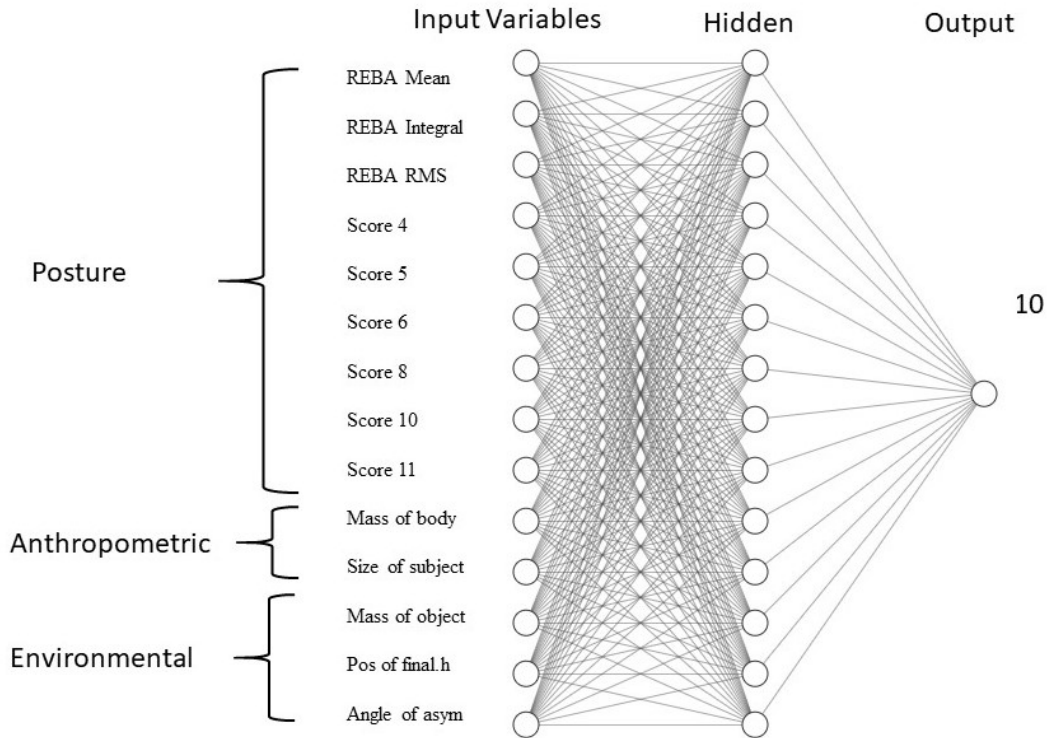
Figure 2. Neural Network structure for the estimation of the discomfort value including one hidden layer and 14 neurons

"k-fold cross validation" (Rodríguez et al. 2010; Fushiki 2011) consisted in using $\frac{1}{k}$ of the data for testing and the remaining $1 - \frac{1}{k}$ for training. After the training step, we computed the validation performance using the mean and standard deviation of the correlation coefficient between predicted and experimental data for each value of k. In this paper, we set $k = 2, 4, 10$, which corresponded to using 50% (resp. 50%), 25% (resp. 75%) and 10% (resp. 90%) of the data for testing (resp. training), as proposed in (Rodríguez et al. 2010).

Bias and variance analysis was used to validate the accuracy of the model, as proposed in (Friedman 1997; Rodríguez et al. 2010). This approach was used not only for the evaluation of the accuracy of the model but also for its ability to generalize to data which were not present in the training data set.

Bias analysis compared the performance of the system with the one obtained by human experts, to estimate a given quantity. If these two performances were close, the neural network was supposed to have similar performance than humans in predicting the reported discomfort CR10. Thus, we needed to quantify the variance τ_h of the answers of the subjects when giving a CR10 score for several times the same situation (same mass, final position, maintaining time). Let us recall that each situation was repeated 3 times by each subject. τ_h was then compared with the error between the predicted and experimental CR10 values, for all the situations. A lower value of this error compared to τ_h would mean that the ANN obtained similar performance than real human during the experiment, when reporting CR10.

The variance analysis aimed at comparing the performance of the neural network estimator when using the test data set compared to the training data set. If RMSEs and correlations using the test data set were similar to those obtained with the training set, it demonstrated the ability of the system to generalize to new data. Reversely, if the RMSE and correlations using the test set were significantly lower than when using the training set, it meant that the system was not able to capture all the possible variance of the relationship, and failed to generalize.

The leave-one-out cross kfold selected subject in the training set (Wong 2015). Such a validation provides relevant information about the ability of the ANN to generalize to new data that have never been used for training. The model was trained on the data of 10 subjects and tested on the data of the missing 11th subject.

Reduction process Previous works (Haj Mahmoud et al. 2020) showed that classical linear data reduction such as Principal Component Analysis did not provide good results. In this paper, we consequently proposed a new approach to iteratively eliminate the variables which do not have a significant influence on discomfort estimation. This greedy process started from 14 available input variables and eliminated one variable at a time, until the performance of the estimation model collapsed.

Hence, for each eliminated variable, we computed the performance of the new best neural network, using the k-fold cross validation, and bias and variance analysis. The new best neural network configuration was used for training, testing and validation processes. We assumed that this greedy method should converge to a global optimal set of variables.

Let R_{14} be the best correlation for a complete 14 variables input set:

$$R_{14} = 1 - \frac{\sum_{i=1}^{N}(y^*(i) - y(i))}{\sum_{i=1}^{N}(y(i))}$$

Where N is the total number of trials in the test data set, $y^*(i)$ is the CR10 collected during the experiments, and $y(i)$ is the estimation found by the neural network.

Then, we performed the same process while eliminating one variable among the 14. We did this for all the 14 variables, and consequently tested 14 possible sets of 13 variables. Let R_{14} be the 4-fold validation (correlation between the predicted and real CR10, when using 75% for training and 25% for testing). We then kept the set of 13 variables whose correlation R_{13} was the closest to R_{14}. This meant that the eliminated variable had the lowest impact on the performance of the neural network. This variable could then be considered as non-significant to estimate discomfort, if the difference between R_{13} and R_{14} was negligible. This process was repeated for all variables and in each step we eliminated the least impacting variable.

The process was stopped when the method could not find a subset of $p < 14$ variables that preserved the performance of the bigger subset $p + 1$. Hence, if a significant decrease of the performance occurred when eliminating a variable, we assumed that the subset $p + 1$ was the minimal subset to correctly estimate CR10.

Results

The correlation between the estimated and experimental CR10 was $R_{14} = 0.775$ for the testing data set, and $R_{14} = 0.822$ for the training data set, when using all the 14 variables (see Table 2). For

$k = 2, 4, 10$, the k-fold cross validation provided us with the following correlations: 0.705, 0.726 and 0.751 respectively.

We also performed a bias and variance analysis, as described before. The RMSE for the testing data set was 1.353.

Once the 14 variables neural network was trained and tested, we iteratively reduced the number of variables. For each subset of variables, we computed the corresponding correlation R_i and $RMSE_i$, with $i < 14$. The results are given in Table 2.

In this Table, n_i stands for the number of neurons in the input layer, and n_h corresponds to the number of neurons in the hidden layer. The number of neurons of the output layer always equaled 1, as it corresponded to the unique CR10 estimation output. Using the above-mentioned steeper slope criterion to stop the iterative process, only 6 variables remained. Indeed, the correlation between predicted and actual discomfort oscillated between 0.763 and 0.782 for all the reduction process steps, but suddenly decreased down to 0.742 and 0.724, when 5 and 4 variables remained respectively.

Table 2. Input variable reduction, starting with the full 14-variables data set. R_{train}, R_{test} stand for the correlation when testing the training or testing data-set respectively. Similarly, $RMSE_{train}$ and $RMSE_{test}$ stand for the RMSE when testing the training and testing data-set respectively. n_i and n_h stand for the number of neurons in the input and hidden layers respectively. Each row indicates the variable which was discarded from the previous subset

Variables	R_{train}	R_{test}	$RMSE_{train}$	$RMSE_{test}$	n_i	n_h
All 14 Variables	0.822	0.775	1.307	1.353	14	14
14 - $REBA_{RMS}$	0.820	0.778	1.295	1.349	13	14
13 - Score 4	0.808	0.763	1.531	1.621	12	14
12 - Score 5	0.821	0.778	1.415	1.520	11	14
11 - Score 10	0.820	0.770	1.352	1.399	10	14
10 - Score 8	0.825	0.781	1.284	1.339	9	14
9 - Score 6	0.812	0.782	1.323	1.363	8	14
8 - Score 11	0.820	0.776	1.467	1.577	7	14
7 - AA	0.826	0.788	1.313	1.326	6	12
6 - $REBA_{Mean}$	0.773	0.742	1.449	1.432	5	12
5 - Pos of final.h	0.754	0.724	1.490	1.518	4	9

For each of the subset of variables, we performed a k-fold cross validation, with $k = 2, 4, 10$, as for the 14-variables data-set. The correlation for the three k values, for all the data subsets is depicted in figure 3.
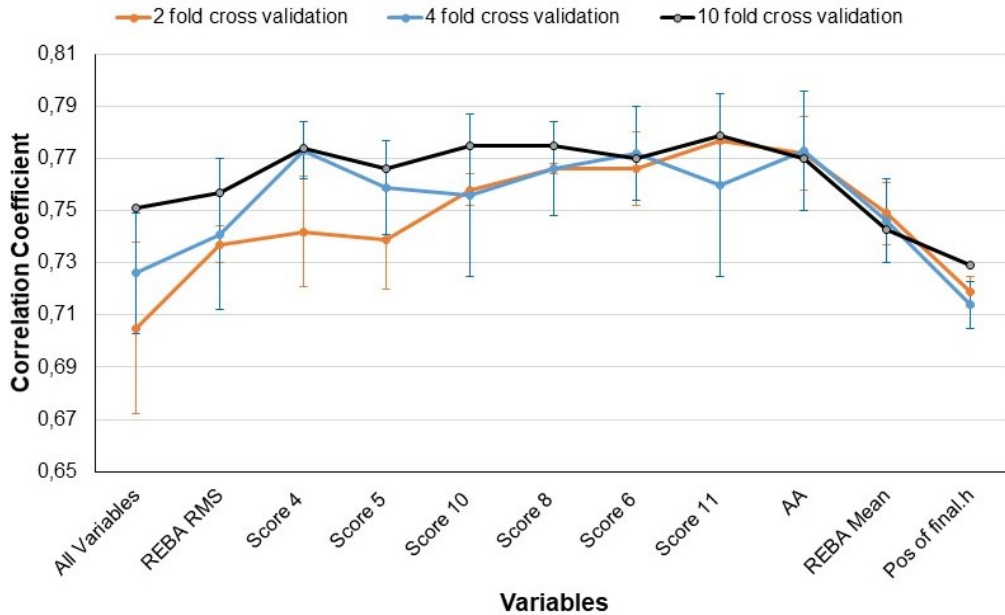
Figure 3. Correlation between predicted and experimental CR10 depending on the tested variable subsets, using the k-fold cross validation, where $k = 2, 4, 10$.

The 6 remaining variables were: mean and integral REBA scores for postural variables, subject mass and stature for anthropometric variables, and final position height and object mass for environmental variables. For this minimal neural network, $R_{test} = 0.826$ and $RMSE_{test} = 1.326$. As for the complete 14-variables set, $RMSE_{test}$ remained lower than the natural standard deviation of reported CR10 scores (1.962). The difference between R_{test} and R_{train} was still small, as for the complete 14-variables set. The k-fold cross validation led to correlations equal to 0.772, 0.773, 0.729 for $k = 2, 4, 10$ respectively.

The results of leave-one-out test showed that the correlation coefficient was equal to 0.709 and the RMSE=1.64 for the complete model with 14 input variables. The correlation coefficient was equal to

0.705 and the RMSE=1.61 for the reduced model with 6 input variables, when applying the leave-one-out cross validation.

Discussion

Contributions

The main contribution of this paper is to analyze and model the interaction between postural, anthropometric and environmental variables to predict reported discomfort in placing and picking tasks. Contrary to most previous works which generally focus on specific variables, such as isolated posture assessment (Chihara et al. 2014; Chryssolouris 2004; Hanson et al. 2003), we used continuous postural score, anthropometric and environmental variables as inputs of the model. The main idea of this work is to avoid assuming any potential interaction between these variables. Consequently, we did not scale the workstation or the carried load according to anthropometric parameters, and we let the machine learning method find potential interactions in the collected data.

Previous works reported weak linear relationship between postural variables (isolated (Chihara et al. 2014) and continuous (Haj Mahmoud et al. 2020) scores) and reported discomfort, especially for light carried loads. Other authors (Li et al. 2009) reported stronger correlation for heavier loads. These results support the idea of a complex relationship between all the variables, not only the mass and the posture. The protocol used in this paper was designed to explore postural variables, but also anthropometric and environmental variables. Both the protocol and the modelling approaches have been designed to avoid any preliminary assumption about interaction between these variables. Hence, the loads and the picking and placing positions were not adapted to the anthropometric data, enabling a larger sampling of the potential interactions between body stature/weight and the task. We used a fully connected multi-layer artificial neural network to model all the possible interactions between the input variables. Our results showed that with 14 variables, the neural network was able to estimate the experimental reported discomfort with a RMSE=1.353 on the testing set, whereas the observed standard deviation of reported CR10 score for all the subjects was τ_h=1.962. It means that the method rated a given situation with a very similar variability

to self-report discomfort. Moreover, the correlation of the results with the testing set - $R_{14} = 0.775$ - reached better performance than previous works (Chihara et al. 2014; Chryssolouris 2004), demonstrating the good performance of the method.

Another main contribution of this work is to search for the minimum set of relevant variables estimating fairly the reported discomfort. As direct linear data reduction approaches, such as PCA and its derivatives, have difficulties to capture the complex interaction between variables (Haj Mahmoud et al. 2020), we proposed a new iterative process to eliminate one variable at a time. This process enabled us to demonstrate that only 6 input variables would be enough to correctly estimate the reported discomfort: mean and integral REBA scores for postural variables, subject mass stature for anthropometric variables, and final position height and object mass for environmental variables. These variables could be measured on site using recent portable motion capture systems (Plantard et al. 2017). This is a promising result to compare the expected discomfort estimated with this model, and the self-reported one, in a more general ergonomic task analysis.

Anthropometric data were included in this minimum set, which was interesting as there was a potential interaction between subject stature and postures. As the starting and final positions were the same for all the subjects, small and tall subjects should have had different postures to achieve the task. Our results support the idea that postural score not only intrinsically reports changes in body stature for the same situation, but also reports differences in locomotor strategies for the same body stature. Further experiments would be necessary to explore this assumption.

Limitations

However, the greedy reduction process could theoretically converge to a local optimal solution and may miss the global optimal solution. In each step of the iterative reduction process, none of the 6 variables that remained at the end, appeared in the 8 (14 total variables minus 6 remaining variables = 8 eliminated variables) less relevant variables. It seems to show that whatever the order in which we eliminate a variable, the 6 remaining variable at the end would be the same, which is promising.

In this work, we used a shallow neural network that was limited by the number of data available for training (constrained by practical limitations of the experimental protocol). This approach was inspired by recent works in other domains (Ma et al. 2017; Steinbach and Altinsoy 2019; Jiménez-Romero et al. 2020) that also used such type of ANN. Any choice in this type of ANN could be discussed, such as the network architecture, the number of layers and neurons per layer, the weights, learning rate, etc. We used some recommendations, such as the choice of the number of weights according to the sample size, or the bias and variance analysis to test the quality of the model. However, several other choices would be possible.Furthermore, without benchmarking with other well-accepted methods (SVM, etc.), it is impossible to judge the true performance of the results. Moreover, machine learning is a very active field of research, especially with the recent development of deep learning approaches, and new network architectures. Although many of these works require a very large amount of data for supervised training, we could explore other types of approaches, such as semi-supervised training, or other types of architectures. In this paper, the main objective was to analyze the interaction between various types of variables to predict discomfort in a specific task. We did not try to innovate in machine learning, and future works would certainly improve our results. However, the results reported in this paper are still very promising, with prediction errors being less than the natural variability when rating CR10 for a subject.

Alwosheel et al. (Alwosheel et al. 2018) consider that the ratio between the number of observations and the number of weights of an ANN should be higher than 10. 11 subjects performed 135 trials with a loss of 3 trials due to missing data, which provided us with a potential set of 1482 input data. As a consequence, the number of weights should be lower than 140 to obtain good performance and to prevent from overfitting. As we did not make any assumption about potential relationships between the input data x, we used a fully-connected ANN, which leads to a high number of weights. For the 14-variables model the ratio between the sample size and the number of weights is 1482/253=6, which is below the expected ratio that led to a ratio of 10. For the 6-variables model, the ratio becomes 1482/61=24. However, the bias and variance analysis showed a limited risk of overfitting, even for the 14-variables model, with a small decrease of performance when applying the model to the training and testing data sets.

In the experimental protocol, the subjects had to repeat 135 picking and placing tasks in an almost 1.5h session. Although it is difficult to completely eliminate the effect of fatigue, we tried to mitigate this problem as much as possible: the order of the 135 trials was randomized for all the subjects, and we introduced as much pauses as we considered as needed. The protocol was tested in preliminary experiments to estimate how much pauses, and the maximum masses and holding times, would be required to avoid fatigue effects in the rating. Because of this limitation of the experimental protocol, it was not possible to test heavier loads, or longer holding times. In future works, it would be interesting to test heavier masses but with shorter holding times, or lighter masses with longer holding times, without introducing fatigue. It could help the model to generalize more on extreme values. In this work, we designed the protocol to eliminate the fatigue factor as much as possible, assuming that each picking and placing task was performed once, isolated. However, in many cases, the tasks are repeated several times, or mixed with other tasks. Hence, to be used on-site, in real situation, it would be important to take repetition and rest times into account.

Pick-and-place tasks are a common action in many industrial processes. However, the results reported in this paper could not be generalized in another context. It seems natural that each task could lead to different interactions between postural, anthropometric and environmental variables. Future works would be needed to better understand the limitations of the current modelling approach for a larger set of tasks. Moreover, it opens fundamental questions, such as the way the relative influence of each variable change according to the task. The cognitive load may also interact with the aforementioned factors, and it may be relevant to include it in the assessment.

This experimentation was led in a laboratory, excluding many of the factors favoring the occurrence of WMSD as in a real industrial context. The relationship between factors would certainly vary in more ecological (industrial) conditions, with noise, extreme temperature, stress, management, health... Therefore, such methods can gain in interest by comparing predicted discomfort to user feedback, to investigate the complementary factors actually responsible for this feedback in an industrial

context. However, preliminary studies are necessary to evaluate the ability of such neural networks approaches to generalize to real situations.

Conclusion

The main contribution of this paper is to analyze the interaction between postural, anthropometric and environmental variables to predict reported discomfort in placing and picking tasks. As this interaction is complex, linear data reduction methods failed to capture the relationship between all these variables. In this paper, we proposed to use a well-known ANN structure to model these complex nonlinear interaction, and designed an innovative iterative data reduction approach to identify the most relevant variables to predict discomfort. The performance of the resulting ANN led to a prediction error which was lower than the measured natural variability in CR10 rating by the subjects. The data reduction methods led to show an interaction between mean and integral REBA scores for postural variables, subject mass and stature for anthropometric variables, and final position height and object mass for environmental variables. This information, and the resulting ANN, could offer new tools to assess plausible discomfort given a picking-and-placing task, and comparing it to actually reported discomfort, in a larger ergonomic assessment of this task. This model can also be used to support ergonomists in designing new workstations and processes, based on digital mock-ups and virtual manikins, at the early stage of the design process.

Future works would help to better understand this complex interaction. Firstly, complementary experiments could help to generalize to a larger set of situations (with heavier masses, longer holding times, or taking repetition into account). Secondly, with the fast development of new machine learning techniques, it would be interesting to evaluate other approaches, and moving from supervised learning to semi-supervised learning to test more complex architectures. Thirdly, the method could be exploited to analyze different tasks, and see how this interaction between variables change depending on the task. It could enable us to design clusters of tasks sharing the same type of interaction between variables.

Key points : Analyze the complex relationship between postural, anthropometric and environmental variables to predict reported discomfort in placing and picking tasks.

As the interaction between these factors and CR10 can't be predicted using linear models, a neural network approach was used.

The performance of the resulting ANN led to a prediction error which was lower than the measured natural variability in CR10 rating by the subjects.

An input data reduction method was proposed, to reduce inputs to the minimum data required to estimate self-reported discomfort with similar accuracy as the neural network fed with all variables.

References

AFNOR (2015). Norme-NF-X35-109. Manutention manuelle de charge pour soulever, déplacer et pousser/tirer.

Alwosheel, A., van Cranenburgh, S., and Chorus, C. G. (2018). Is your dataset big enough? sample size requirements when using artificial neural networks for discrete choice analysis. *Journal of Choice Modelling*, 28:167–182.

Ando, S., Ono, Y., Shimaoka, M., Hiruta, S., Hattori, Y., Hori, F., and Takeuchi, Y. (2000). Associations of self estimated workloads with musculoskeletal symptoms among hospital nurses. *Occupational and Environmental Medicine*, 57(3):211–216.

Bhatnager, V., Drury, C. G., and Schiro, S. G. (1985). Posture, postural discomfort, and performance. *Human Factors*, 27(2):189–199.

Borg, G. (1998). Borg's perceived exertion and pain scales. *Human Kinetics*, (July 1998):104 vii.

Chihara, T., Hoshi, S., and Seo, A. (2014). 1B2-2 Fundamental Study on Formulation of Total Perceived Discomfort for the Entire Body in the Sagittal Plane based on Joint Moment. *The Japanese Journal of Ergonomics*, 50(Supplement):S140–S141.

Chryssolouris, D. Pappas, M. K. V. (2004). Posture Based Discomfort Modelling using NEURAL NET. *Manufacturing, Modelling, Management and Control*, pages 19–23.

Da Costa, B. R. and Vieira, E. R. (2010). Risk factors for work-related musculoskeletal disorders: A systematic review of recent longitudinal studies. *American Journal of Industrial Medicine*, 53(3):285–323.

Das, B. and Gangopadhyay, S. (2011). An ergonomics evaluation of posture related discomfort and occupational health problems among rice farmers. *Occupational Ergonomics*, 10(1-2):25–38.

Friedman, J. H. (1997). On bias, variance, 0/1-loss, and the curse-of-dimensionality. *Data Mining and Knowledge Discovery*, 1(1):55–77.

Fushiki, T. (2011). Estimation of prediction error by using K-fold cross-validation. *Statistics and Computing*, 21(2):137–146.

Georgiou, T., Liu, Y., Chen, W., and Lew, M. (2019). A survey of traditional and deep learning-based feature descriptors for high dimensional data in computer vision. *International Journal of Multimedia Information Retrieval*, 9(3):135–170.

Haj Mahmoud, O., Pontonnier, C., Dumont, G., Poli, S., and Multon, F. (2020). Posture assessment and subjective scale agreement in picking tasks with low masses. In *Advances in Intelligent Systems and Computing*, pages 32–38. Springer International Publishing.

Hanson, L., Wienholt, W., and Sperling, L. (2003). A control handling comfort model based on fuzzy logics. *International Journal of Industrial Ergonomics*, 31(2):87–100.

Hignett, S. and McAtamney, L. (2000). Rapid Entire Body Assessment (REBA). *Applied Ergonomics*, 31(2):201–205.

Jiménez-Romero, F. J., Guijo-Rubio, D., Lara-Raya, F. R., Ruiz-González, A., and Hervás-Martínez, C. (2020). Validation of artificial neural networks to model the acoustic behaviour of induction motors. *Applied Acoustics*, 166:1–11.

Kee, D. and Karwowski, W. (2001a). LUBA: An assessment technique for postural loading on the upper body based on joint motion discomfort and maximum holding time. *Applied Ergonomics*, 32(4):357–366.

Kee, D. and Karwowski, W. (2001b). The boundaries for joint angles of isocomfort for sitting and standing males based on perceived comfort of static joint postures. *Ergonomics*, 44(6):614–648.

Li, K. W., feng Yu, R., Gao, Y., Maikala, R. V., and Tsai, H.-H. (2009). Physiological and perceptual responses in male chinese workers performing combined manual materials handling tasks. *International Journal of Industrial Ergonomics*, 39(2):422–427.

Liao, M. H. and Drury, C. G. (2000). Posture, discomfort and performance in a vdt task. *Ergonomics*, 43(3):345–359.

Ma, C., Chen, C., Liu, Q., Gao, H., Li, Q., Gao, H., and Shen, Y. (2017). Sound Quality Evaluation of the Interior Noise of Pure Electric Vehicle Based on Neural Network Model. *IEEE Transactions on Industrial Electronics*, 64(12):9442–9450.

Markus Hofman and Ralf Klinkenberg (2014). *RapidMiner Data Mining Use Cases and Buisness Analytics Applications*. Taylor & Francis e-Library.

McAtamney, L. and Corlett, N. (1993). RULA: A survey method for the investigation of work-related upper limb disorders. *Applied Ergonomics*, 24(2):91–99.

National Research Council and Institute of Medicine (2001). *Musculoskeletal Disorders and the Workplace: Low Back and Upper Extremities*.

Okuyucu, H., Kurt, A., and Arcaklioglu, E. (2007). Artificial neural network application to the friction stir welding of aluminum plates. *Materials and Design*, 28(1):78–84.

Plantard, P., Shum, H. P., Le Pierres, A. S., and Multon, F. (2017). Validation of an ergonomic assessment method using Kinect data in real workplace conditions. *Applied Ergonomics*, 65:562–569.

Rodríguez, J. D., Pérez, A., and Lozano, J. A. (2010). Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):569–575.

Ruck, D. W., Rogers, S. K., and Kabrisky, M. (1990). Feature selection using a multilayer perceptron. *Journal of Neural Network Computing*, 2(2):40–48.

Steinbach, L. and Altinsoy, M. E. (2019). Prediction of annoyance evaluations of electric vehicle noise by using artificial neural networks. *Applied Acoustics*, 145:149–158.

Walters, T. R., Putz-Anderson, V., and Garg, A. (1994). cdc_5434_DS1 NIOSH. *DHHS(NIOSH) Publication No.94-110*.

Wong, T. T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, 48(9):2839–2846.

Wu, G. and Cavanagh, P. R. (1995). ISB recommendations for standardization in the reporting of kinematic data. *Journal of Biomechanics*, 28(10):1257–1261.

Zhang, Q., Yang, L. T., Chen, Z., and Li, P. (2018). A survey on deep learning for big data. *Information Fusion*,

42:146–157.