

Contributing Components of Metabolic Energy Models to Metabolic Cost Estimations in Gait

Markus Gambietz, Marlies Nitschke, Jörg Miehling, Anne D. Koelewijn

Abstract—Objective: As metabolic cost is a primary factor influencing humans' gait, we want to deepen our understanding of metabolic energy expenditure models. Therefore, this paper identifies the parameters and input variables, such as muscle or joint states, that contribute to accurate metabolic cost estimations. **Methods:** We explored the parameters of four metabolic energy expenditure models in a Monte Carlo sensitivity analysis. Then, we analysed the model parameters by their calculated sensitivity indices, physiological context, and the resulting metabolic rates during the gait cycle. The parameter combination with the highest accuracy in the Monte Carlo simulations represented a quasi-optimized model. In the second step, we investigated the importance of input parameters and variables by analysing the accuracy of neural networks trained with different input features. **Results:** Power-related parameters were most influential in the sensitivity analysis and the neural network-based feature selection. We observed that the quasi-optimized models produced negative metabolic rates, contradicting muscle physiology. Neural network-based models showed promising abilities but have been unable to match the accuracy of traditional metabolic energy expenditure models. **Conclusion:** We showed that power-related metabolic energy expenditure model parameters and inputs are most influential during gait. Furthermore, our results suggest that neural network-based metabolic energy expenditure models are viable. However, bigger datasets are required to achieve better accuracy. **Significance:** As there is a need for more accurate metabolic energy expenditure models, we explored which musculoskeletal parameters are essential when developing a model to estimate metabolic energy.

Index Terms—Biomechanics, Deep Learning, Gait, Metabolic Cost, Sensitivity Analysis.

NOMENCLATURE

ANN	Artificial Neural Network
CMC	Coefficient of Multiple Correlation
KIMR15	MEE Model by Kim et al.
KSstat	Kolmogorov–Smirnov Statistic
LICH05	MEE Model by Lichtwark et al.
LOO	Leave-One-Out

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1483 – Project-ID 442419336, EmpkinS.

Markus Gambietz, Marlies Nitschke, and Anne D. Koelewijn are with the Machine Learning and Data Analytics Lab, Department Artificial Intelligence in Biomedical Engineering (AIBE), Faculty of Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91052 Erlangen, Germany (email: markus.gambietz@fau.de)

Jörg Miehling is with the Chair of Engineering Design, Department of Mechanical Engineering, Faculty of Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany

MARG68	MEE Model by Margaria
MC	Monte Carlo
MEE	Metabolic Energy Expenditure
MINE97	MEE Model by Minetti et al.
RMC	Repeated Measures Correlation
RMSE	Root Mean Square Error

I. INTRODUCTION

Over the last decades, improvements in hard- and software, as well as an improved understanding of gait, have enabled researchers to design smart active wearable devices, such as prostheses and exoskeletons. Those devices can improve gait and, thereby, participation in society for persons with a gait disability. One of the ways these devices improve participation in society is by reducing the energy consumption or metabolic cost of walking. Since the middle of the last century, experimental research has shown that people choose many gait variables, such as speed, cadence and stride length, to reduce the metabolic cost of gait [1], [2]. Furthermore, it was recently shown that even in minutes to hours, persons would adjust their gait to perform this optimization [3], [4]. Therefore, there is a need for accurate assessment of metabolic energy consumption.

Metabolic energy expenditure (MEE) can be assessed computationally via MEE models and experimentally, the latter of which is more commonly used. These experiments use direct or indirect calorimetry. In direct calorimetry, the heat produced by the human body is measured directly, for example, in a heat chamber [5]. The system setup and operation are complex and costly [5]. Therefore, indirect calorimetry is used more frequently. In indirect calorimetry, the energy expenditure is determined by measuring the amount of inhaled and exhaled oxygen and carbon dioxide [6]. Oxygen consumption is directly related to energy expenditure when the performed exercise is aerobic. However, also indirect calorimetry has limitations. First, it requires about 3 minutes to adjust to the level of the oxygen inhalation required for the exercise [7]. Secondly, breathing is variable; consequently, to get an accurate estimate of energy expenditure based on the average over several breaths, a more extended measurement period is required. Therefore, participants must perform the exercise of interest for at least 5 minutes, though longer periods could improve accuracy. Hence, indirect calorimetry estimates the averaged MEE over the measurement period and does not provide instantaneous MEE estimates at a specific point in time. Furthermore, direct and indirect calorimetry can only

provide full-body energy expenditure, including exercise and the expenditure of the brain, the gastrointestinal system, and other organs. Therefore, it is not straightforward to calculate the "net metabolic cost" or the metabolic cost of the activity itself, and the contribution of different muscles cannot be individually assessed.

Instead, MEE models could be used to estimate instantaneous energy expenditure at the muscle level [8], [9]. MEE models calculate energy expenditure based on kinetic and kinematic movement variables. Margaria [10] was the first to describe a relationship between eccentric and concentric muscle contraction based on muscle efficiency during shortening and lengthening. Further MEE models tried to form new approaches for the computational determination of MEE. They were based on, for example, human gait in two dimensions (2D) [11], three dimensions (3D) [12], or the contraction of the sartorius muscle of a frog [13]. Depending on the approach, MEE models vary in complexity and contain between two and fifteen empirical parameters, while the amount of parameters required to represent muscle physiology accurately remains unclear.

So far, comparative studies have focused on the accuracy and correlation of the MEE models compared to measured metabolic rates. Miller et al. [8] compared five MEE models at self-chosen walking speed and cadence. In the study by Koelewijn et al. [9], the models from [8], and two additional MEE models were tested for their responses in altered environments, e.g., different walking speeds and inclines. Both studies reported an underestimation of metabolic cost in gait [8], [9]. A recent study aimed to estimate changes in energy expenditure in older people due to changes in walking patterns using MEE models. However, the MEE models could only explain about one-third of the measured changes [14]. Therefore, the accuracy of MEE models should be improved to predict changes in the metabolic rate correctly.

Furthermore, the use of MEE models in simulations of human gait has proven to be successful in replicating natural walking patterns by minimizing energy-related objectives [15] with the use of both simple [16] and complex musculoskeletal models [8]. Muscle-space MEE models have been used as the objective [17], [18] for muscle-driven simulations. For torque-driven simulations, a joint-space MEE model [12] could be used as the objective. In such gait simulations, the objective function's complexity significantly impacts the computational time [17]. Hence, from a simulation standpoint, there is a motivation to use accurate MEE models with lower complexity to enhance computational efficiency.

In addition to the challenges associated with the accuracy of MEE models, there are other important considerations to address. In biomechanical processing, musculoskeletal models often rely on cadaveric measurements of elderly individuals as a basis [19]. However, when scaling these generic musculoskeletal models, crucial muscle parameters required for accurate calculation of metabolic costs are not personalized [20]. Moreover, muscle force estimation is resource intensive and depends on the specific optimization problem formulation [21]. Consequently, the calculation of metabolic costs is already based on inherently inaccurate data. An alternative

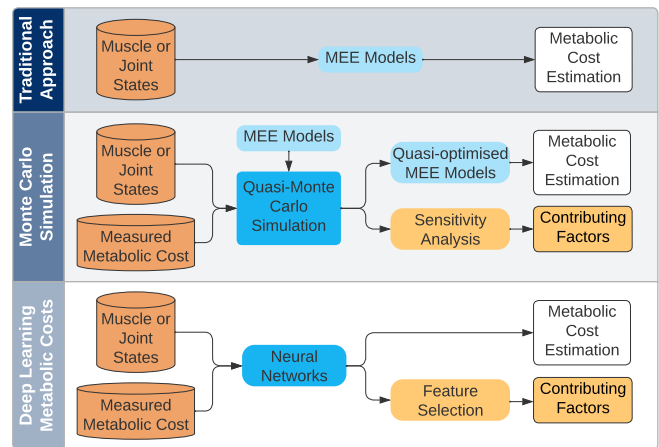


Fig. 1: Overview of the approach in this paper. We identified the contributing factors in current MEE models locally via a sensitivity analysis. Furthermore, we searched for contributing components globally by performing a feature selection using neural networks.

approach to mitigate some of the limitations at the muscle level is to estimate metabolic costs at the joint level [12]. Furthermore, metabolic cost models cannot be fully validated since energy expenditure cannot be measured at a muscle level during gait.

In this study, we aim to deepen our understanding of MEE estimations by investigating the importance of different model parameters both locally and globally (see Figure 1). We investigate the sensitivity of model parameters of the different MEE models to understand which parameters are essential when developing more accurate MEE models. First, we undertook a local sensitivity analysis by performing a Monte Carlo simulation on the empirical model parameters to determine their influence quantitatively. Furthermore, we identified the MEE models with the highest accuracy in the Monte Carlo simulation and investigated how the MEE estimations of these models differ from the original version. In the second step, we searched for the globally most optimal feature set. Therefore, we used neural networks as purely statistical representations for metabolic models and trained them with different input modalities.

II. METHODS

A. Data set

We have used all available trials of the publicly available data set of Koelewijn et al. [22]. It comprises six female and six male participants, averaging 24 ± 5 years in age, 173 ± 8 cm in height and a weight of 70 ± 12 kg. Gait of each participant was recorded at two velocities (0.8 m s^{-1} , 1.3 m s^{-1}) and three inclines (-8% , 0% , 8%).

The data set contains processed data consisting of joint angles, joint velocities, and joint moments, of the hip, knee and ankle in the sagittal plane, as well as muscle activations, stimulations, and lengths, for eight muscles in each leg. The data processing was described and performed by Koelewijn et

al. [9]. Koelewijn et al. estimated joint angles, joint velocities, and joint moments using a 2D link-segment leg model. Then, they used a dynamic optimization procedure [21] to estimate muscle activations, stimulations, and length. Furthermore, the average metabolic cost for each trial, measured through indirect calorimetry, was reported [22]. This data can be used directly as inputs for the MEE models described in section II-B.

B. Metabolic energy expenditure models

In our sensitivity analysis, we compared and analyzed all models from Koelewijn et al.'s study [9]. We present the MEE models MARG68 [10], MINE97 [11], LICH05 [23], and KIMR15 [12] in the manuscript, while the MEE models by Bhargava et al. [13], Houdijk et al. [24], and Umberger et al. [25] are added in the supporting document, since these models are conceptionally similar to LICH05. MARG68, MINE97, and LICH05 use muscle states, while KIMR15 uses joint states to estimate the instantaneous metabolic rate \dot{E} [W]. The metabolic cost C_{calc} can be derived from the metabolic rate [9]:

$$C_{calc} = \frac{1}{Tmv} \int_{t=0}^T \sum_{i=0}^{n_{mus}} \dot{E}_i dt, \quad (1)$$

where T is the duration of the motion, m the body mass, v the velocity, and n_{mus} the number of muscles i for which the metabolic rate \dot{E} is summed.

MARG68: The model by Margaria [10] is based on the observation of differences in MEE between muscle lengthening and shortening. Margaria estimated that the shortening efficiency η_s is 25%, with the lengthening efficiency η_l being 120% in relation to the power \dot{w} , depending on the sign of the fibre velocity v_{ce} :

$$\dot{E} = \begin{cases} \frac{\dot{w}}{\eta_s} & v_{ce} < 0 \\ -\frac{\dot{w}}{\eta_l} & v_{ce} \geq 0. \end{cases} \quad (2)$$

The power \dot{w} is the product of muscle velocity v_{ce} and muscle contractile element force F_{ce} , which itself is the product of maximum isometric force F_{max} , activation a , and the Hill-type force-length $f_{FL}(\tilde{l}_{ce})$ and force-velocity $f_{FV}(\tilde{v}_{ce})$ relationships.

MINE97: The model of Minetti and Anderson [11] estimates the metabolic rate \dot{E} from the shortening velocity based on an empirical relationship:

$$\dot{E} = av_{max}F_{max}\phi(v_{ce}), \quad (3)$$

where a is the muscle activation, v_{max} the maximum muscle fibre velocity, and F_{max} the maximum isometric force. ϕ is a relationship between fibre velocity and energy expenditure, with $\tilde{v}_{ce} = v_{ce}/v_{max}$ being the relation of the normed muscle fibre velocity \tilde{v}_{ce} to the maximum muscle fibre velocity v_{max} :

$$\phi(v_{ce}) = \frac{p_{mi,1} + p_{mi,2}\tilde{v}_{ce} + p_{mi,3}\tilde{v}_{ce}^2}{p_{mi,4} + p_{mi,5}\tilde{v}_{ce} + p_{mi,6}\tilde{v}_{ce}^2 + p_{mi,7}\tilde{v}_{ce}^3}. \quad (4)$$

The parameters values $p_{mi,1...7}$ of the original model are [0.054, 0.506, 2.46, 1, -1.13, 12.8, -1.64].

LICH05: In the model by Lichtwark et al. [23], the total energy rate \dot{E} is calculated through the sum of power \dot{w} and heat rate \dot{h} produced due to muscle shortening and lengthening as well as maintenance of the activation. In this context, the parameters $p_{l,1...8}$ represent the unnamed empirical parameters that are present in the model. The maintenance heat rate is dependent on a decay function γ and the fibre velocity where shortening is defined as positive:

$$\dot{h}_M = \gamma \frac{\tilde{v}_{max}}{G^2} \begin{cases} 1 & v_{ce} > 0 \\ (p_{l,2} + (1 - p_{l,2})e^{-p_{l,3}\tilde{v}_{ce}(f_{FV}(\tilde{v}_{ce})-1)}) & v_{ce} \leq 0, \end{cases} \quad (5)$$

where $f_{FV}(\tilde{v}_{ce})$ is the Hill-type force-velocity characteristic, G is the curvature of $f_{FV}(\tilde{v}_{ce})$, and γ is defined as:

$$\gamma = p_{l,4}e^{-p_{l,5}t_{stim}} + p_{l,6}e^{-p_{l,7}t_{stim}}, \quad (6)$$

where t_{stim} is the duration for which the activation of a muscle has been higher than a threshold t_{act} . The shortening-lengthening heat rate is defined as:

$$\dot{h}_{SL} = \begin{cases} \frac{\tilde{v}_{ce}}{G} & v_{ce} > 0 \\ -p_{l,8}f_{FV}(\tilde{v}_{ce})\tilde{v}_{ce} & v_{ce} \leq 0. \end{cases} \quad (7)$$

A modified version [26] of the original LICH05 model [23] was used, where the overall heat rate is defined as:

$$\dot{h} = aF_{max} \left(p_{l,1}\dot{h}_M + f_{FL}(\tilde{l}_{ce}) \left((1 - p_{l,1})\dot{h}_M + \dot{h}_{SL} \right) \right). \quad (8)$$

The empirical parameters of LICH05 were the following: $G = 4$, $t_{act} = 0.1$, and $p_{l,1...8} = [0.3, 0.3, 7, 0.8, 0.72, 0.175, 0.022, 0.5]$.

KIMR15: The model created by Kim et al. [12] is the only joint-space MEE model and calculates the metabolic rate based on joint angles q , joint velocities \dot{q} , and joint moments M :

$$\dot{E} = \dot{h}_M|\dot{q}|_{max}|M| + \dot{h}_{SL}|M\dot{q}| + \dot{q}_{cc}(M\dot{q})_{max} + M\dot{q}, \quad (9)$$

where $\dot{h}_M = 0.054$, $\dot{h}_{SL,s} = 0.283$ for shortening and $\dot{h}_{SL,l} = 1.423$ for lengthening, respectively, and $\dot{q}_{cc} = 0.004$ is the co-contraction heat rate.

C. Quasi-Monte Carlo sensitivity analysis

We conducted a sensitivity analysis to determine which parameters in the MEE models have a large or small influence on the calculated metabolic cost in gait. For this purpose, we used a quasi-Monte Carlo (MC) simulation by drawing a large number of random MEE model parameter sets, assessing the error and correlation to the measured metabolic cost for every randomised parameter set. The random parameter sets include all empirical parameters in the respective MEE models. An overview of the respective search spaces per parameter is listed in the supporting document.

In each iteration, each gait trial's metabolic cost was calculated using a quasi-randomly parametrised MEE model. The MEE model parameters were sampled using Saltelli's extension of the Sobol sequence, as implemented in PyTorch

[27]–[29]. The quasi-random Sobol sequences enable the generation of low-discrepancy sequences, which cover the space of possible parameter combinations more evenly than pseudo-random methods. As a metric, the root-mean-squared error (RMSE) and the repeated measures correlation (RMC) [30] between the calculated and measured metabolic cost were determined. The RMSE is an indicator for the absolute accuracy of the MEE models, while the RMC is an indicator for intra-subject correlation. Furthermore, the coefficient of multiple correlation (CMC) [31] was used to assess the shape similarity of metabolic rate curves between the MEE models.

After sampling the model parameters and calculating the corresponding RMSE, the sensitivity indices S_i were computed via the Kolmogorov-Smirnoff statistic (KSstat) [32] with the scipy library's "kstest" [33]. Extensive testing showed that a combination of 10^5 model evaluations with a behavioural set size of 10^2 was suitable for both exploring the parameter space and calculating the KSstat. These sensitivity indices can be used to quantitatively rank the individual parameters according to their contribution to the metabolic cost estimation. We enhanced the understanding of simplest models, MARG68 and MINE97, by incorporating visual representations to illustrate the Monte-Carlo simulation. This was not possible for the other models due to the high number of parameters.

Furthermore, the model parameter sets with the lowest RMSE values in the MC simulation resulted in quasi-optimised MEE models since those estimated the measured metabolic cost best. We performed a subject-wise leave-one-out (LOO) cross-validation to report the RMSE and RMC achieved with the quasi-optimised MEE models.

D. Deep MEE Models

We trained artificial neural networks (ANNs) to estimate MEE using different feature combinations of muscle and joint parameters and variables as input to make statements about the necessity of these features. Furthermore, we determined whether ANNs can outperform the classical MEE models on our data set. Like the MEE models shown in section II-B, we trained the ANNs with combinations of muscle or joint states as an input vector to return metabolic rate per muscle or joint at each time point.

ANNs are a purely statistical model that does not take prior knowledge of the biological system into account, allowing us to search for feature combinations globally. As shown in Figure 2, every combination of the input features used in the different MEE models was tested to assess which feature combinations, in both joint and muscle space, are crucial for MEE predictions. In the muscle space, the eight input features are normalised fibre length \tilde{l}_{ce} , normalised fibre velocity \tilde{v}_{ce} , muscle activation a , maximum isometric force F_{max} , fast-twitch fibre percentage r_{ft} , muscle width w_h , optimal fibre length $l_{ce,opt}$, and muscle stimulation e . The joint-space ANNs received feature combinations of joint angles q , velocities \dot{q} , accelerations \ddot{q} , and moments M . Each input was scaled to a minimum / maximum value between 0 and 1.

The ANNs estimate the metabolic rate of each joint or muscle at each time point, as the traditional metabolic models

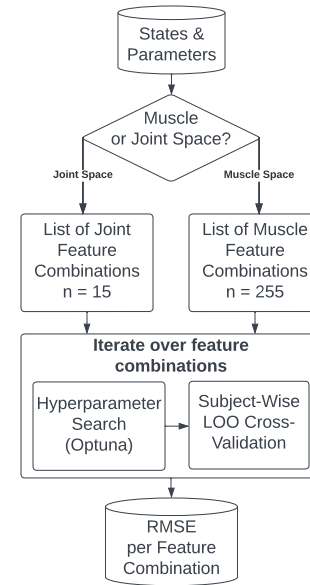


Fig. 2: Deep-Learning MEE model pipeline.

do. After the forward pass, the metabolic rate is summed over all muscles or joints and then averaged over the gait cycle. Then, the loss is formed by comparing the estimated average metabolic rate to the gait cycle's average metabolic rate computed from the measured metabolic cost. Therefore, our training approach can be classified as inexact supervised learning since the metabolic rate for each muscle and time point is indirectly supervised [34]. We used this approach to maintain the ability of MEE models to estimate individual muscle or joint metabolic rates and to be able to investigate the metabolic rate over time.

All models were set up as dense networks in PyTorch [29], [35], where a combination of the mean-squared error loss function and Adam as the optimiser [36] proved most reliable. ReLU and Leaky ReLU activation functions were used [37]–[39]. Hyperparameter tuning was done using Optuna's TPE-Sampler [40], with tunable hyperparameters being the number of layers, layer sizes, batch size, learning rate, and weight decay.

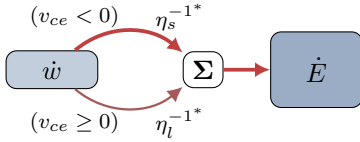
The feature combinations were evaluated by comparing their respective RMSE after subject-wise LOO cross validation. The training set of each fold, which contains the trials from the remaining eleven subjects, was randomly split into 70 % training and 30 % test set. For each feature combination, hyperparameter tuning was performed during the first cross-validation step. The networks were trained on a cluster equipped with NVIDIA RTX 2080s.

III. RESULTS

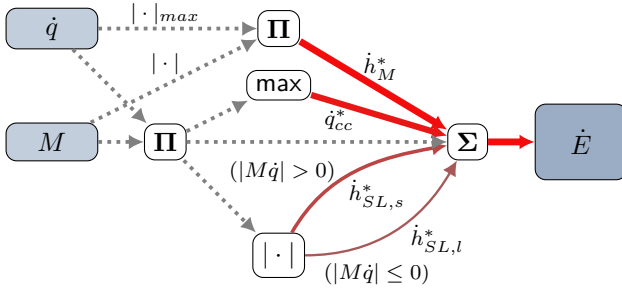
A. Sensitivity analysis on MEE model parameters

The identification of the most sensitive parameters is model-dependent (see Figure 3). In the case of MINE97 and KIMR15, the parameters associated with force ($p_{mi,1}$ in MINE97 and \dot{h}_m in KIMR15) exhibit the highest sensitivity. Conversely, in MARG68 and LICH05, the parameters associated with shortening exhibit the highest sensitivity indices. In KIMR15, the

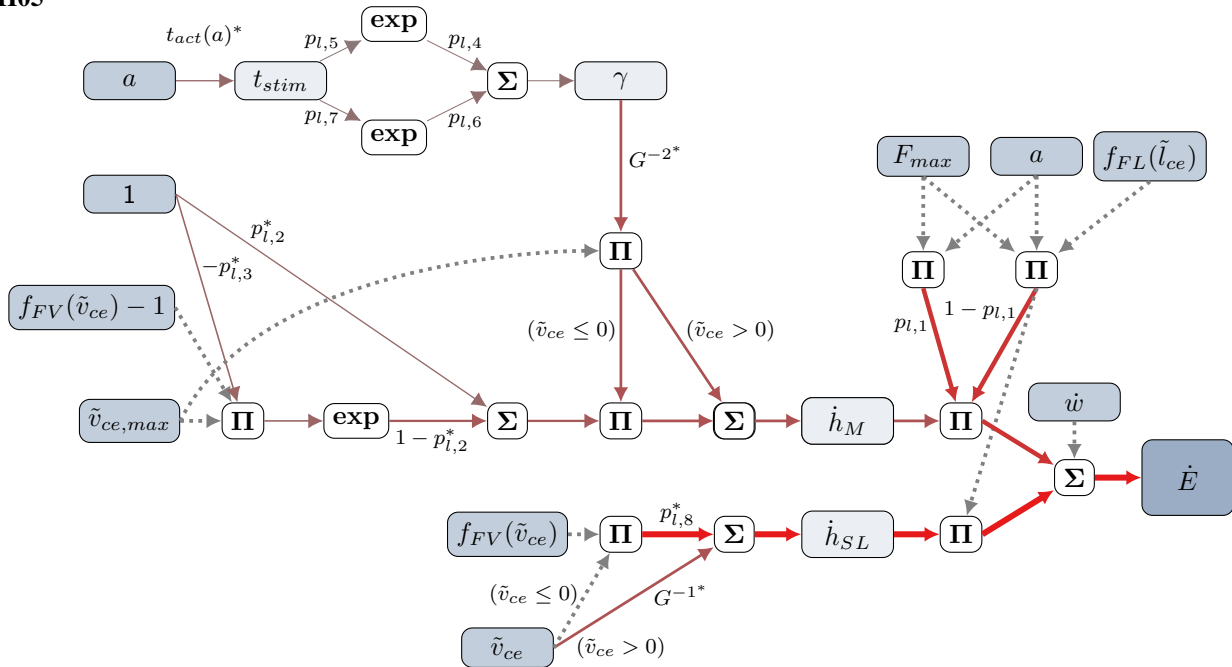
MARG68



KIMR15



LICH05



MINE97

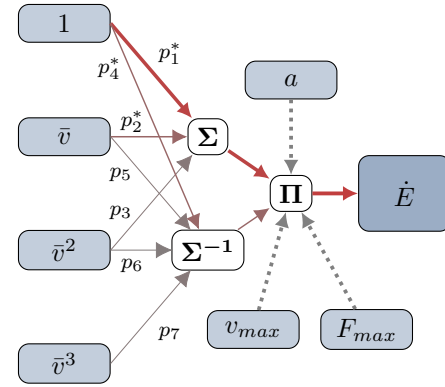


Fig. 3: These flowcharts illustrate the computation steps of the respective MEE models and the sensitivity associated with the parameters at each computation step. The thicker red arrows indicate higher sensitivity indices, while the thinner gray arrows indicate lower sensitivity indices. Parameters with significant sensitivity indices ($p < 0.05$) are marked with an asterisk. The gray dotted arrows represent input propagation without modifications. When multiple arrows end at the same node, the brighter one is used for continuation. The metabolic rate \dot{E} is presented in dark blue boxes, while inputs are shown in blue, and intermediate results (refer to Chapter II-B) are shown in light blue boxes.

shortening parameter also outranks the lengthening parameter in sensitivity. A complete list of sensitivity indices can be found in the supporting document. However, note that the indices cannot be compared between the MEE models because each is based on a separate sensitivity analysis.

After performing the quasi-Monte Carlo simulation, the quasi-optimised MEE models better fit the measured metabolic

cost than the original MEE models since the RMSE improved on average by $44 \pm 13\%$ and the RMC increased slightly (see Table I). Figure 4 shows exemplary metabolic rate curves the all muscle-space MEE models. Between the curves of the original MEE models, a CMC of 0.84 was observed. The quasi-optimised MEE models show a CMC of 0.92 between each other, meaning the similarity between the curves is more

prominent in the quasi-optimised MEE models, and they are all optimised towards the same shape. In the following, the individual models are described and pictured.

TABLE I: Overview of all repeated measures correlations (RMC) and root mean square errors (RMSE) for the original and quasi-optimised MEE models. The RMC and RMSE were obtained from the Monte Carlo simulation using a leave-one-out cross-validation.

Model	RMC	Original	Quasi-Optimised MEE	
		RMSE [J kg ⁻¹ m ⁻¹]	RMC	RMSE [J kg ⁻¹ m ⁻¹]
MARG68	0.90	1.19	0.92	0.79
MINE97	0.95	1.00	0.93	0.73
LICH05	0.96	1.06	0.96	0.68
KIMR15	0.91	1.98	0.89	1.15

MARG68: In MARG68, the KSstat sensitivity index S_i for shortening is 0.50 and higher than the one for lengthening with 0.30. As MARG68 contains only two parameters, the RMSE for each parameter combination can be visualised in a 2D-plane (see Figure 5). A minimum RMSE was found at a shortening efficiency η_s of 25.7 % and a lengthening efficiency η_l of -228 %, corresponding to 44 % energy restoration.

MINE97: Similar to MARG68, the MINE97 model also estimates the metabolic rate as a function of fibre velocity. In MINE97's polynomial ϕ -function, the lower-order parameters ($p_{mi,1,4}$) generally have the highest sensitivity indices. Figure 6 shows the best fitting ϕ -functions from the MC simulation and compares it to the original MINE97 model. During shortening, the best fits estimate a higher metabolic rate, while during lengthening, the best fits predict negative metabolic rates.

LICH05: As visible in Figure 3, LICH05 is the most complex MEE model in our study, containing a shortening-lengthening as well as a maintenance heat rate. The results show that the highest sensitivity index belongs to parameter $p_{l,8}$, associated with muscle shortening. The second highest sensitivity index belongs to parameter $p_{l,1}$, regulating the impact of Hill-type force-length relationship on the maintenance heat rate. On the other hand, the parameters of the γ -function of the LICH05 model, which are related to the activation time dependency, have the lowest sensitivity indices.

KIMR15: In KIMR15, KSstats for maintenance heat rate $S_i = 0.91$ and co-contraction heat rate $S_i = 0.83$ dominate. The sensitivity of the shortening-lengthening heat rate S_i is 0.47 for shortening and 0.31 for lengthening.

B. Deep MEE Models

We found that for muscle and joint space, the minimum achieved RMSEs were 0.87 J kg⁻¹ m⁻¹ and 0.91 J kg⁻¹ m⁻¹, respectively. Therefore, none of the ANNs was able to achieve a better RMSE than the quasi-optimised MEE models. In the muscle space, the feature combination with the lowest RMSE was contractile element velocity \tilde{v}_{ce} , muscle stimulation e , activation a , maximum isometric force F_{max} and the fast-twitch fibre percentage r_{ft} . The 20 best results used muscle activation, stimulation or both, and used fibre velocity \tilde{v}_{ce} . Further features that were encountered among most of the

top results were maximum isometric force F_{max} and the fast twitch fibre percentage r_{ft} .

The combination of joint space features with the lowest RMSE was the joint moment M , velocity \dot{q} and acceleration \ddot{q} . Omitting \ddot{q} increased the RMSE by 7.2 %. Every other change in feature selection saw an RMSE increase of at least 16.8 %. The RMSE achieved by the quasi-optimised MEE version of KIMR15 was 26 % higher than the best-performing joint-space ANN.

In summary, including power-related features played a crucial role in enabling ANNs to achieve accurate estimations of metabolic costs. However, the force-related features (M , a , e) alone already contribute to the moderate accuracies observed in Figure 7. In contrast, the velocity-related features (\dot{q} , \tilde{v}_{ce}) show improved performance only when combined with force-related features. A full list of feature combinations and their achieved RMSEs can be found in the supporting document.

IV. DISCUSSION

This study examined the influence of MEE model parameters on the metabolic cost estimations and thus which muscle or joint states and parameters are crucial for the metabolic cost estimations of gait. The three muscle-space MEE models featured between two and ten empirical parameters. Table I shows that the quasi-optimised version of the more complex LICH05 model (see Figure 3) only slightly outperforms the quasi-optimised versions of the simpler models, similar to the results reported in [9]. We concluded that for estimating metabolic costs in (healthy) walking, simpler metabolic models are sufficient. As the CMC of 0.92 between the quasi-optimised MEE models indicates, the quasi-optimised MEE models converged towards a mutual MEE estimation for each muscle over time. Figure 4 shows the curve shape similarity, especially between LICH05 and MINE97, suggesting that their formulation is well suited for fitting gait data.

The local sensitivity analysis conducted in both muscle and joint space revealed that the parameters with the highest sensitivity were those related to work. Unsurprisingly, this finding was globally supported by the ANNs used in the study. Furthermore, force-related parameters proved valuable in traditional MEE models and ANNs, even without including velocity-related parameters. Therefore, the suggested separate maintenance heat rate by LICH05 and KIMR15 is valid. The four traditional MEE models might benefit from including fast-twitch fibre ratios, as this parameter was present in the highest-ranking ANN, suggesting its potential contribution. Fast-twitch fibre ratios have been employed in several other MEE models before [13], [24], [25].

A key finding was that the MEE curves of the quasi-optimised MEE models tend to show higher peak positive metabolic rates. In most cases, the original MEE models underestimated metabolic costs. Therefore, the higher metabolic rates are not surprising. Furthermore, the quasi-optimised MEE models also showed negative metabolic rates during lengthening (see Fig. 4). As shortening and lengthening exist during the gait cycle, the timing of energy consumption is shifted from more extreme negative MEE rates towards

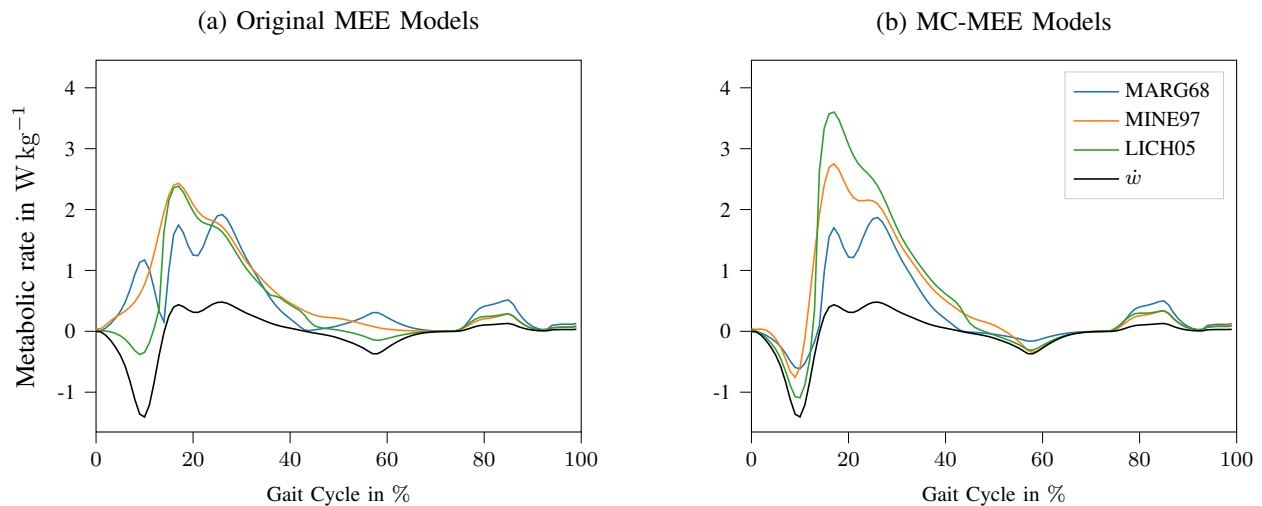


Fig. 4: A representative participant's vasti muscular metabolic rate according to the muscle-space MEE models in level walking. Left: Original MEE models. Right: Best-performing quasi-optimised MEE models from the MC simulation. Power \dot{w} of the muscle annotated in black.

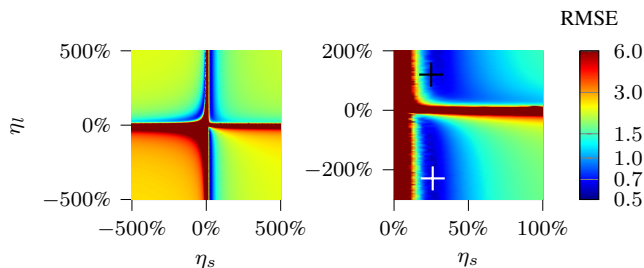


Fig. 5: Heatmap for MARG68 in $[\text{J kg}^{-1} \text{m}^{-1}]$ showing the relationship between the shortening η_s and lengthening η_l efficiency and the RMSE. On the right zoomed-in heatmap, the original MARG68 model is pointed out with a black cross, and the quasi-optimised MARG68 parameters are annotated with a white cross.

peak positive MEE rates. As seen in Figures 5 and 6, the quasi-optimised version of the MEE models would have a negative efficiency, or ϕ -function in MINE97's case, during lengthening. As depicted in Figure 4, the models fail to estimate sufficient heat to compensate for the energy delivered to the muscle, which is not thermodynamically reasonable. True negative metabolic rates would refer to the synthesis of adenosine triphosphate, which is impossible [8], [41]. It is plausible that modeling errors have contributed to this behavior. In particular, muscle states were estimated via dynamic optimisation using Hill's muscle model. This muscle models is based on measurements of isotonic contractions near resting length and only accurate for such conditions [42]. We found a possible explanation in the winding filament theory [43], which suggests better capabilities for muscles to store elastic energy, thus explaining the negative MEE rates we found, as energy stored during lengthening and released during shortening. To eliminate the possibility of this observation being due to the dataset used, we repeated the Monte Carlo

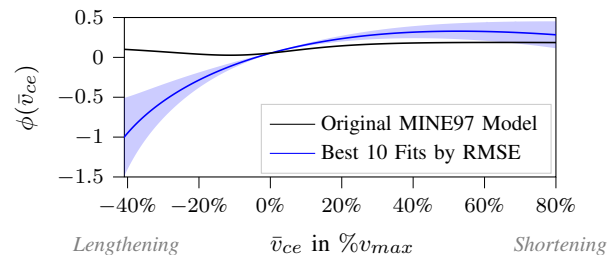


Fig. 6: MINE97's ϕ -function. Mean and standard deviation of the ten best fits from the MC simulation and the original MINE97 model. Minimum and maximum values of \bar{v}_{ce} were chosen according to the values occurring in the dataset.

simulation on a dataset by Theunissen *et al.* [44]. Here, we also found a negative lengthening efficiency for the quasi-optimised MEE models.

In this work, the ANNs could not outperform the quasi-optimised MEE models. We attribute the reason to the inexact supervision combined with the relatively small training data set. The utilization of inexact supervision was preferred to enable the ANNs to mimic the traditional MEE model, which estimates metabolic rate using muscle/joint states at each time point. In the future, however, ANNs could potentially outperform other MEE models. For example, we trained a temporal convolutional network on marker trajectories and ground reaction forces that reached 20% lower error compared to any of the tested feature combinations, performing slightly worse than the quasi-optimised LICH05 model. Recurrent ANNs such as temporal convolutional networks, LSTMs or transformer ANNs can replicate time-dependent muscle properties, as attempted in the activation-time-dependency in BHAR04 and LICH05. Furthermore, using raw data such as marker trajectories and ground reaction forces can prevent error accumulation during processing. Potential error sources can be inconsistencies in inverse dynamics and improper

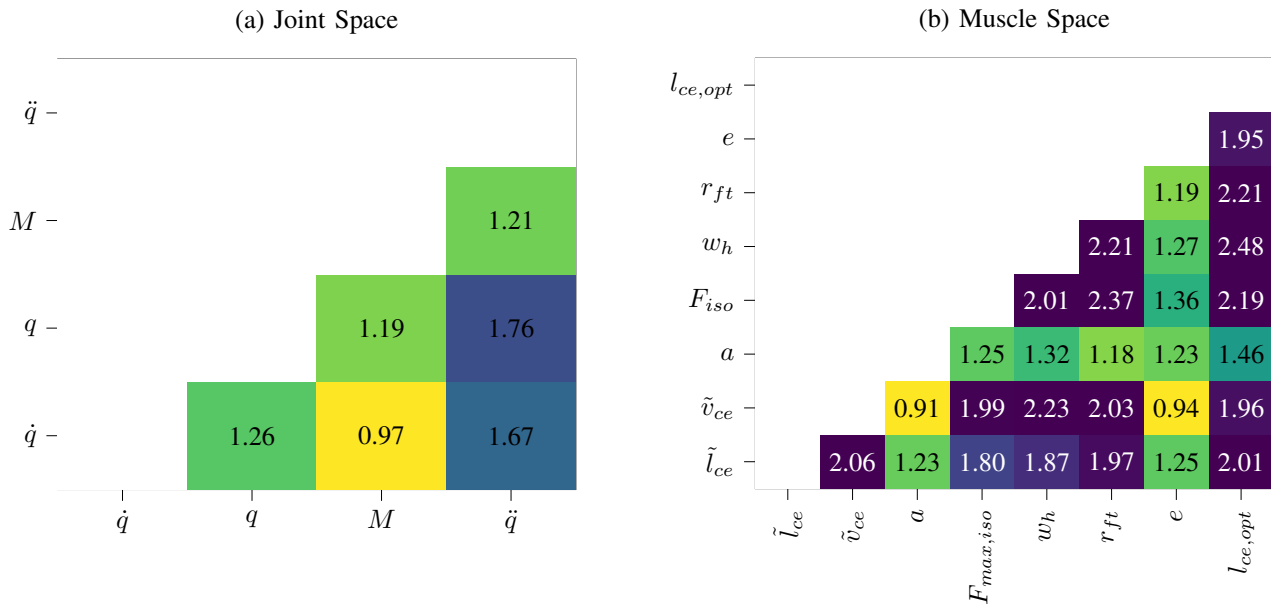


Fig. 7: Comparison of RMSE [$J\ kg^{-1}\ m^{-1}$] for joint (a) and muscle space (b) across various feature combinations. The RMSE corresponds to an ANN precisely trained only those two features. The color scale represents the error magnitude, with yellow indicating lower error and blue representing higher error.

personalisation of musculoskeletal models.

To get first insights into the robustness of the quasi-optimised MEE and ANN models, which were trained on the data of Koelwijn et al. [22], we tested the models with other publicly available datasets of Theunissen et al. [44] and Pimentel et al. [14]. On both datasets, the quasi-optimised MEE models saw a slight in- or decrease in RMSE over their original counterparts. ANNs did not generally outperform the quasi-optimised MEE models on these datasets either. This was expected since our training dataset was small and not necessarily representative for the conditions measured in those studies. It is important to note that the inter-MEE model discrepancy is generally less pronounced than the inter-dataset discrepancy in terms of metabolic cost estimation accuracy, likely due to different processing methods and settings.

In addition to more accurate MEE models, further advancements in biomechanical processing are necessary to improve metabolic cost estimations. Reconstruction through inverse kinematics and inverse dynamics with general musculoskeletal models leads to dynamic inconsistencies, where residual forces remain at the pelvis [45]. Processing differences between different models are large [46]. To reduce both, a unified pipeline providing accurate model personalisation would be helpful. For example, the current scaling method implemented in OpenSim [20] does not personalise any muscle-related parameters such as maximum isometric forces, moment arms and fibre type percentages. However, the actual values of these muscle-related parameters vary vastly among the population [47], [48]. Most MEE models base their estimations on muscle-related parameters, implying the need for personalisation of these factors as well.

Based on our outcomes, we have defined some recommendations about which MEE model to choose for specific

use cases. For instance, simpler MEE models can be used to calculate the metabolic cost of walking, as we have shown that they are similarly accurate as more complex models. However, it should be further tested if they can be used as objectives for walking simulations since computer optimizations might find unrealistic solutions when models are too simple. Furthermore, in other applications, especially those with more isometric contractions, we expect that more complex models are required. For example, MARG68 assigns no metabolic costs to isometric contractions and would likely provide inaccurate results. Similarly, the parameters of MINE97 are based on walking experiments. Nonetheless, we argue that the future development of MEE models should aim for agnosticism regarding specific movement types, while also focusing on minimal computational complexity to allow its application across a broad spectrum of scenarios.

V. CONCLUSION

We addressed the question which parameters and muscle/joint states are crucial in metabolic cost estimations. In the local sensitivity analysis, we found that power-related parameters play the biggest role in metabolic cost estimations. Increasing the complexity of MEE models and the number of inputs does not necessarily increase the model's estimation accuracy.

When exploring deep-learning as an alternative to classical MEE models, we could also globally confirm the non-necessity of many variables by training ANNs, because including some of the features only delivered marginally better or even worse results. Using ANNs, did not achieve better accuracy than the quasi-optimised MEE models. Because we found that the inter-MEE model discrepancy is generally less pronounced than the inter-dataset discrepancy in terms of metabolic cost estimation

accuracy, we argue for unified processing methods as a basis for validation of future work in MEE modelling.

ACKNOWLEDGMENT

We thank Dr Dario Zanca and Dr An Nguyen for their advice on the deep learning approach. We thank Dr Jason Franz and Dr Richard Pimentel from the University of North Carolina at Chapel Hill for sharing their data and the general discussion about metabolic energy consumption.

REFERENCES

- [1] H. J. Ralston, "Energy-speed relation and optimal speed during level walking," *Internationale Zeitschrift für angewandte Physiologie, einschliesslich Arbeitsphysiologie*, vol. 1958, no. 4, pp. 277–283, 1958.
- [2] M. Y. Zarrugh *et al.*, "Optimization of energy expenditure during level walking," *European journal of applied physiology and occupational physiology*, vol. 33, no. 4, pp. 293–306, 1974.
- [3] J. C. Selinger *et al.*, "Humans can continuously optimize energetic cost during walking," *Current Biology*, vol. 25, no. 18, pp. 2452–2456, 2015.
- [4] J. M. Finley *et al.*, "Learning to be economical: the energy cost of walking tracks motor adaptation," *The Journal of Physiology*, vol. 591, no. 4, pp. 1081–1095, 2013.
- [5] J. A. Levine, "Measurement of energy expenditure," *Public Health Nutrition*, vol. 8, no. 7a, p. 1123–1132, 2005.
- [6] E. Ferrannini, "The theoretical bases of indirect calorimetry: A review," *Metabolism*, vol. 37, no. 3, pp. 287–301, 1988.
- [7] J. C. Selinger and J. M. Donelan, "Estimating instantaneous energetic cost during non-steady-state gait," *Journal of Applied Physiology*, vol. 117, no. 11, pp. 1406–1415, 2014.
- [8] R. H. Miller, "A comparison of muscle energy models for simulating human walking in three dimensions," *Journal of biomechanics*, vol. 47, no. 6, pp. 1373–1381, 2014.
- [9] A. D. Koelewijn *et al.*, "Metabolic cost calculations of gait using musculoskeletal energy models, a comparison study," *PloS one*, vol. 14, no. 9, p. e0222037, 2019.
- [10] R. Margaria, "Positive and negative work performances and their efficiencies in human locomotion," *Internationale Zeitschrift für angewandte Physiologie, einschliesslich Arbeitsphysiologie*, vol. 25, no. 4, pp. 339–351, 1968.
- [11] A. E. Minetti and R. M. Alexander, "A theory of metabolic costs for bipedal gaits," *Journal of theoretical biology*, vol. 186, no. 4, pp. 467–476, 1997.
- [12] J. H. Kim and D. Roberts, "A joint-space numerical model of metabolic energy expenditure for human multibody dynamic system," *International journal for numerical methods in biomedical engineering*, vol. 31, no. 9, p. e02721, 2015.
- [13] L. J. Bhargava *et al.*, "A phenomenological model for estimating metabolic energy consumption in muscle contraction," *Journal of biomechanics*, vol. 37, no. 1, pp. 81–88, 2004.
- [14] R. E. Pimentel *et al.*, "Muscle metabolic energy costs while modifying propulsive force generation during walking," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 0, no. 0, pp. 1–14, 2021.
- [15] J. E. Bertram and A. Ruina, "Multiple walking speed–frequency relations are predicted by constrained optimization," *Journal of Theoretical Biology*, vol. 209, no. 4, pp. 445–453, 2001.
- [16] M. Srinivasan and A. Ruina, "Computer optimization of a minimal biped model discovers walking and running," *Nature*, vol. 439, no. 7072, p. 72–75, 2006, cited by: 417; All Open Access, Green Open Access.
- [17] A. D. Koelewijn *et al.*, "A metabolic energy expenditure model with a continuous first derivative and its application to predictive simulations of gait," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 21, no. 8, pp. 521–531, Jun. 2018. [Online]. Available: <https://www.tandfonline.com/doi/full/10.1080/10255842.2018.1490954>
- [18] A. Falisse *et al.*, "Rapid predictive simulations with complex musculoskeletal models suggest that diverse healthy and pathological human gaits can emerge from similar control strategies," *J. R. Soc. Interface.*, vol. 16, no. 157, p. 20190402, Aug. 2019. [Online]. Available: <https://royalsocietypublishing.org/doi/10.1098/rsif.2019.0402>
- [19] V. Zatsiorsky, "Kinetics of human motion," *Human Kinetics*, pp. 79–110, 2002.
- [20] S. L. Delp *et al.*, "Opensim: open-source software to create and analyze dynamic simulations of movement," *IEEE transactions on bio-medical engineering*, vol. 54, no. 11, pp. 1940–1950, 2007.
- [21] F. de Groot *et al.*, "Evaluation of direct collocation optimal control problem formulations for solving the muscle redundancy problem," *Annals of biomedical engineering*, vol. 44, no. 10, pp. 2922–2936, 2016.
- [22] A. D. Koelewijn *et al.*, "Dataset for metabolic cost calculations of gait using musculoskeletal energy models, a comparison study," 2018.
- [23] G. A. Lichtwark and A. M. Wilson, "A modified hill muscle model that predicts muscle power output and efficiency during sinusoidal length changes," *The Journal of experimental biology*, vol. 208, no. Pt 15, pp. 2831–2843, 2005.
- [24] H. Houdijk *et al.*, "Evaluation of a hill based muscle model for the energy cost and efficiency of muscular contraction," *Journal of biomechanics*, vol. 39, no. 3, pp. 536–543, 2006.
- [25] B. R. Umberger *et al.*, "A model of human muscle energy expenditure," *Computer methods in biomechanics and biomedical engineering*, vol. 6, no. 2, pp. 99–111, 2003.
- [26] G. A. Lichtwark and A. M. Wilson, "Is achilles tendon compliance optimised for maximum muscle efficiency during locomotion?" *Journal of biomechanics*, vol. 40, no. 8, pp. 1768–1775, 2007.
- [27] I. Sobol, "Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates," *Mathematics and Computers in Simulation*, vol. 55, no. 1-3, pp. 271–280, 2001.
- [28] A. Saltelli *et al.*, "Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index," *Computer Physics Communications*, vol. 181, no. 2, pp. 259–270, 2010.
- [29] A. Paszke *et al.*, "Pytorch: An imperative style, high-performance deep learning library," 2019.
- [30] J. Z. Bakdash and L. R. Marusch, "Repeated measures correlation," *Frontiers in psychology*, vol. 8, p. 456, 2017.
- [31] A. Ferrari *et al.*, "A new formulation of the coefficient of multiple correlation to assess the similarity of waveforms measured synchronously by different motion analysis protocols," *Gait & posture*, vol. 31, no. 4, pp. 540–542, 2010.
- [32] F. Pianosi *et al.*, "Sensitivity analysis of environmental models: A systematic review with practical workflow," *Environmental Modelling & Software*, vol. 79, pp. 214–232, 2016.
- [33] P. Virtanen *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.
- [34] Z.-H. Zhou, "A brief introduction to weakly supervised learning," *National Science Review*, vol. 5, no. 1, pp. 44–53, 08 2017.
- [35] S. Bai *et al.*, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv:1803.01271*, 2018.
- [36] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017.
- [37] V. Nair and G. Hinton, "Rectified linear units improve restricted boltzmann machines," *Proceedings of ICML*, vol. 27, 2010.
- [38] I. Goodfellow *et al.*, *Deep learning*. Cambridge, Massachusetts and London, England: MIT Press, 2016.
- [39] B. Xu *et al.*, "Empirical evaluation of rectified activations in convolutional network," 2015.
- [40] T. Akiba *et al.*, "Optuna: A next-generation hyperparameter optimization framework," 2019.
- [41] R. C. Woledge *et al.*, "Energy storage during stretch of active single fibres," *Advances in experimental medicine and biology*, vol. 538, pp. 627–33; discussion 634, 2003.
- [42] D. A. Winter, *Biomechanics and motor control of human movement*, 4th ed. Hoboken, N.J: Wiley, 2009.
- [43] K. C. Nishikawa *et al.*, "Is titin a 'winding filament'? a new twist on muscle contraction," *Proceedings. Biological sciences*, vol. 279, no. 1730, pp. 981–990, 2012.
- [44] K. Theunissen *et al.*, "Dataset of energetics and biomechanics of self-paced and fixed speed treadmill walking at multiple speeds," *Data in Brief*, vol. 41, p. 107915, 2022.
- [45] J. L. Hicks *et al.*, "Is My Model Good Enough? Best Practices for Verification and Validation of Musculoskeletal Models and Simulations of Movement," *Journal of Biomechanical Engineering*, vol. 137, no. 2, 02 2015.
- [46] A. Falisse *et al.*, "Opensim versus human body model: A comparison study for the lower limbs during gait," *Journal of Applied Biomechanics*, vol. 34, no. 6, pp. 496 – 502, 2018.
- [47] S. K. Powers and E. T. Howley, *Exercise physiology: Theory and application to fitness and performance*, 8th ed., ser. Connect, learn, succeed. New York, NY: McGraw-Hill, 2012.
- [48] B. Heitmann, "Body fat in the adult danish population aged 35-65 years: an epidemiological study," *International journal of obesity*, vol. 15, no. 8, p. 535–545, August 1991.