Title: Predicting steady state metabolic power in cerebral palsy, stroke, and the elderly during walking with and without assistive devices

Authors: Karl Harshe[1], Benjamin C. Conner[2], Zachary F. Lerner[1,2]

Author affiliations:

[1]Department of Mechanical Engineering, Northern Arizona University, Flagstaff, AZ, USA

[2]College of Medicine – Phoenix, University of Arizona, Phoenix, AZ, USA

Corresponding Author:

Zachary F. Lerner, PhD
Associate Professor
Mechanical Engineering Department
Northern Arizona University
Flagstaff AZ, 86001
Phone: 928-523-1787
Fax: 928-523-2300
Email: Zachary.Lerner@nau.edu
Web: http://biomech.nau.edu

Word count: 7426

Abstract

Purpose. Individuals with walking impairment, such as those with cerebral palsy, often face challenges in leading physically active lives due to the high energy cost of movement. Assistive devices like powered exoskeletons aim to alleviate this burden and improve mobility. Traditionally, optimizing the effectiveness of such devices has relied on time-consuming laboratory-based measurements of energy expenditure, which may not be feasible for some patient populations. To address this, our study aimed to enhance the state of the art predictive model for estimating steady state metabolic rate from two minute walking trials to include individuals with and without walking disabilities and for a variety of terrains and wearable device conditions. **Methods.** Using over 200 walking trials collected from eight prior exoskeleton-related studies, we trained a simple linear machine learning model to predict metabolic power at steady state based on condition-specific factors, such as whether the trial was conducted on a treadmill (level or incline) or outdoors, as well as demographic information, such as the participant's weight or presence of walking impairment, and two minutes of metabolic data. **Results.** We demonstrated the ability to predict steady state metabolic rate to within an accuracy of 4.71 ± 2.7% for varying walking conditions, such a differing terrain, patient populations. **Conclusion.** This work seeks to unlock the use of in-the-loop optimization of wearable assistive devices in individuals with limited walking capacity. A freely available MATLAB application allows other researchers to easily apply our model.

Key Terms: Cerebral palsy, metabolic steady state, machine learning, exoskeleton

Competing Interests: None of the authors have a competing interest to declare.

1. Introduction

High energetic cost of movement is a significant barrier to a physically active lifestyle for those with deficits in neuromuscular control due to brain injury, such as in cerebral palsy (CP) [1] and stroke [2]. This is particularly detrimental in pediatric populations, like those with CP, where physical activity is an essential stimulus for healthy development [3]. The underlying mechanisms of these neuromuscular deficits appears to be multifaceted and often patient-dependent, but are linked to impaired coordination that leads to ineffective kinetics and kinematics [4], [5], [6]. The resulting increase in metabolic demand can limit functional performance and make fundamental tasks, such as activities of daily living, difficult to complete [7]. This high energetic demand with movement is the basis for several established and evolving interventions for individuals with CP that are aimed at improving mobility by reducing its metabolic burden, such as passive bracing [8], [9], [10] and powered exoskeleton devices [11], [12]. Similarly, independent mobility is closely linked with quality of life in the elderly [13], which has also led to the exploration of assistive exoskeletons in this population [14], [15], [16]. The clinical acceptance of these assistive technology interventions is often predicated on their ability to reduce energy expenditure and increase movement capacity in patient populations.

Demonstrating reduced energy expenditure with assistive devices has often relied on laboratory-based measurements of whole-body energy expenditure via gold standard

techniques such as indirect calorimetry. Indirect calorimetry works by measuring the inspired and expired gases that correlate with metabolic activity within an individual's musculature to estimate energy usage [17]. By measuring changes in metabolic power from one activity (e.g., quiet sitting) to another (e.g., walking), a net difference can be calculated (i.e., net metabolic power) that represents the energetic cost of the latter activity relative to the former. This net metabolic power is a measure of how much energy a person's body expends on any given activity or task. A significant disadvantage to this gold standard approach is the amount of time required to measure metabolic power. Indirect calorimetry is inherently noisy and there is a delay in how inspired and expired gases change with increased muscular metabolism due to mitochondrial dynamics and gas transport [18], requiring individuals to partake in potentially strenuous walking activity for extended periods of time (i.e., 5-10 minutes) until a "steady state" is reached and an average of measured values can be taken to reduce noise. This time requirement can be challenging when working with patient populations where extended activity under multiple testing conditions is not feasible.

The time requirement for measuring steady state metabolic power is also limiting for a leading machine learning approach used in optimizing assistive devices via human-in-the-loop optimization [19]. This methodology uses an iterative process to determine optimal design or control settings of an assistive device to maximally reduce a specific metric, such as metabolic power. The process consists of walking with a specific assistive setting (e.g., level of assistive torque) while steady state metabolic power is measured or estimated, after which an optimization algorithm chooses a new assistive

setting and metabolic power is once again measured or estimated, repeating until the algorithm can converge on the most optimal assistive settings. The outcomes of this optimization process have been promising for establishing user-specific control settings [19], [20], [21], [22]. However, measuring steady state metabolic power for each iteration of a human-in-the-loop optimization scheme may not be possible for many individuals with disabilities because they may be unable to walk for the time required for convergence. For these reasons, there is motivation to predict steady state metabolic rate from short walking bouts prior to reaching steady state.

Various methods have demonstrated promise for accurately predicting steady state metabolic rate during walking in unimpaired individuals. These methods include predictions based on anthropometrics and walking speed [23], first order physiological signals [24], time derivatives of physiological signals [25], and the mechanics of walking [26]. A prominent approach is based on a predictive model utilizing the first two minutes of walking metabolic data to approximate the metabolic response as a first-order linear system, essentially creating a two-minute weighted time series prediction [22]. This method represents the current State of the Art (SoA) for human-in-the-loop optimization of lower limb wearable devices. However, this method has not been evaluated in, or developed for, individuals with neuromuscular disorders. This is of particular importance because factors such as atypical muscle mass and physiology [27], deficits in neuromuscular coordination, and abnormal walking kinematics [28] can have a significant influence on energy expenditure, making predictive equations developed from unimpaired populations much less accurate for those with impairments where

these factors are affected [29]. We expect that such a model would capture the biomechanical causes (e.g., spasticity, sarcopenia, etc.) of differences in metabolic power across patient populations during walking with and without assistive devices. This void hampers the ability of researchers to individually optimize wearable devices for people with walking disabilities, particularly when their disability limits the amount of time they can ambulate.

The overarching objective of this study was to establish a model to quickly and accurately estimate steady state metabolic power, with a similar accuracy (4-6% error) and time requirements (2 minutes) as the current SoA, in individuals with impaired ambulation under a variety of walking conditions, including on different terrains (e.g., level or incline), or with wearable assistive devices (e.g., exoskeletons) (Fig. 1). Our primary goal was to validate and expand on, if necessary, the weighted time series prediction model to accurately predict steady state metabolic rate from short, transient walking trials of two minutes or less in individuals with CP and other walking impairments. Given the complex interactions between movement impairment, assistive device use, terrain, and metabolic power, we hypothesized that adding demographic and condition-specific parameters to the current weighted time series prediction model would result in a more accurate prediction of metabolic power when compared to the weighted time series prediction model alone. ‘Demographic information’ included information about the person and any groups to which they belonged, specifically – age, height, weight, presence of a CP diagnosis or stroke history, being in an elderly population or being unimpaired. ‘Condition-specific parameters’ refer to how the

experiment was conducted: Treadmill or overground terrain, surface grade, the presence of an exoskeleton or ankle-foot orthotic, and walking speed. This allowed us to compare the accuracy of our new predictive model with the existing weighted time series prediction model across several clinical populations (CP, stroke, and the elderly). This exploratory analysis sought to evaluate the potential for a single model that could be applied across different assistive device research topics and patient populations. Therefore, we also included a single stroke participant in this investigation. *We did, however, wish to see how a general model would function when applied to minimal training data, and one stroke participant was included.* We also had a secondary goal to evaluate whether the prediction accuracy of our model would increase if we captured the rate at which metabolic power increased at the start of each trial (i.e., “ramp rate”: the slope of the best fit line calculated on the first two minutes of metabolic data recorded from each walking trial).

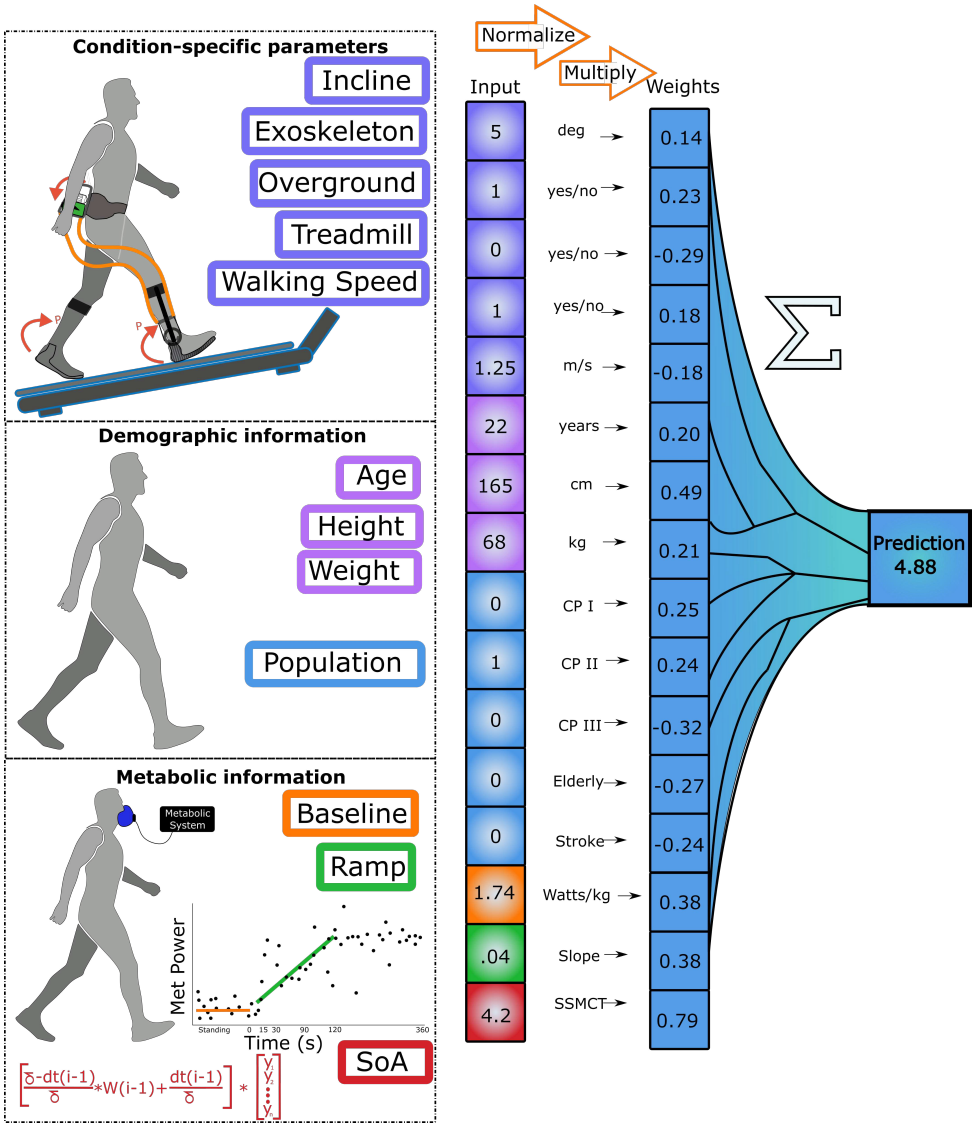

Figure 1. Metrics used to create the dataset included personal information, trial details and metabolic time series data. Boolean values (input format depicted as 0 or 1) marked each of the 5 populations, as well as specific trial conditions which could be combinations of treadmill, overground, shod, exoskeleton assistance, zero-torque and incline. The remaining variables required a continuous input data format. Metabolic data were initially collected from both VO_2 and VCO_2, but were combined into a single metabolic power metric and normalized to body mass. SoA model from [22].

2. Materials and Methods

We implemented a retrospective analysis of walking data to generate a flexible linear model for predicting steady state metabolic power in those with and without neuromuscular impairment and under various walking conditions. The data for this retrospective analysis were acquired from both prior and concurrent investigations at

Northern Arizona University [12], [14], [30], [31], [32], [33]. The studies were approved by the Institutional Review Board (IRB) of Northern Arizona University under protocols #986744-1:45. All participants, or their parents/guardians if under 18 years old, provided informed consent, or verbal assent, if applicable, before participation. Each trial was represented by a vector of variables with data format depicted in Fig. 1. Participants included individuals with cerebral palsy, stroke, elderly individuals, and those without walking impairments. All data were collected over the period spanning 2018 to 2024. The dataset employed in this study comprised three primary data types: 1). Metabolic data obtained from a metabolic sampling system (either Parvo (TrueOne, USA) or Cosmed (K-5, USA)), 2); basic participant characteristics such as height, age, weight, clinical population, and impairment level for CP (Gross Motor Function Classification Score); and 3) experimental metrics including walking speed, presence of an exoskeleton or walking aid, and terrain type. CP is divided into five Gross Motor Function Classifications (I-V), where CP I is the least impaired and CP V has the greatest impairment to motor control [34]. Consequently, the highest energetic costs associated with ambulation are found in individuals with the CP III-V [35].The simplified distributions of these demographics and the associated testing conditions are provided in Table 1.

Table 1. Condition Distribution

	Shod	Zero-torque	Assistance	AFO
CP I (n = 8)	48	30	96	0
CP II (n = 8)	12	15	54	0
CP III (n = 3)	12	12	39	3
Elderly (n = 11)	8	2	19	0
Stroke (n = 1)	2	2	6	0
Unimpaired (n = 13)	0	66	168	0

Table 1. Distribution of the total utilized observations (inclusive of all three interopolation methods) across populations and walking aids. Shod represents walking in shoes only, zero torque represents walking with the exoskeleton worn but not imparting any significant additional forces, Assistance indicates that the device was providing stance-phase torque assistance while walking, and AFO indicates that the participant wore rigid or semi-rigid ankle-foot-orthoses and no exoskeleton. CP I-III represents Gross Motor Function Classificaiton Score Levels I-III respectively.

a. Model Design Objectives

We considered two factors for establishing prediction accuracy targets for benchmarking our models. First, due to the inherent variability in metabolic measurements, we sought to predict steady state metabolic power to within the 95% confidence interval (CI) of the measurement average. The 95% CI for the data used in our analysis (i.e., model labels) was 0.3-5.2%, with an average error rate of 2.5%. This finding was consistent with a prior report of metabolic system accuracy ranging from 1.7 ± 0.9% to 6.8 ± 6.5% measurement error [36]. Second, we also reviewed the accuracy of weighted time series prediction when used in the context of human-in-the-loop optimization. Predicting metabolic power to 4.3% accuracy, on average, with a maximum reported error of 15.5%, has been demonstrated to be adequate for use in human-in-the-loop optimization of exoskeleton assistance in unimpaired individuals [37]. Therefore, to serve as an effective tool for in-the-loop optimization, our model should

achieve a comparable accuracy of 4-6%, which was very similar to the target established from our 95% CI calculation. To maximize the potential for our model to facilitate the optimization of assistive devices in individuals with limited walking endurance, we sought to assess model accuracy at the current best practice (120 second trial duration). Additionally, we sought to make a freely available MATLAB application that would allow other researchers to easily input participant and walking condition characteristics and apply our predictive model(s) in near real-time.

b. Linear Model Building

This study investigated the predictive capabilities of supervised learning, a category of machine learning using labeled data to predict outcomes, combined with stepwise linear regression models in estimating steady state metabolic power [38]. We chose linear models and not more complex implementations for interpretability and ease of use [39]. Linear models assign a weight to each input variable and the sum of all weighted variables is the prediction. Stepwise linear regression uses statistical methods to iteratively select the most significant predictor variable to build a model that best fits the data, and it does not include features that do not significantly improve model accuracy. Using this architecture, we were able to examine how including additional parameters (such as the demographic and condition-specific parameters mentioned above) impacted our linear model's accuracy.

MATLAB (2023b) was used to implement a stepwise linear regression and identify statistically significant features. We maintained a significance cutoff of $p < 0.05$ for

features to be added to the model, as well as requiring the p-value to remain below this level to be retained in the final model. We then used only the features which were identified as significant to train, validate, and test our PyTorch implementation of linear regression; PyTorch is a freely available machine learning library for Python [40], [41]. The linear model used input data that were identified as significant from the linear regression and used the validated adaptive moment estimation (ADAM) optimizer contained in the PyTorch framework [42].

Data were initially pooled and shuffled, then ten-fold cross-validation was employed to identify the optimal training duration, and the validation error for each of the ten folds (i.e., set or partition of our dataset) was retained for accuracy assessment. Best practices for cross-validation were employed, including tracking validation error for each fold, and identifying overtraining prior to the cessation of training on that fold [43]. Additionally, validation errors were recorded for all folds and the average validation error of the folds was used to identify the optimal number of training epochs prior to testing. Interval scaling was used in conjunction with the scikit-learn standard scalar to ensure that scaled observations would remain in the positive domain and on the same order of magnitude as all other features to achieve model convergence [44]. Test data, comprised of samples unseen by the model during training, were used to evaluate the final performance of the model. These data were identified as a subset of the data in each fold that used the repeated measurements interpolation method (an in-depth explanation of the validation and testing splitting methodology is available in the supplementary materials). Finally, the loss function used during optimization was the

minimization of mean squared error [45]. To identify optimal training times for each fold, at the end of every epoch of training (i.e., complete pass of the training dataset through the algorithm), the validation data were evaluated with the mean squared error loss function, and recorded along with the accuracy of the training data. Once all epochs were completed, the epoch corresponding to the minimum validation error was identified and that duration was used to train on that fold and predict on the test set. This resulted in ten test subset errors, which were used to compare model accuracies. All fold model validation data were subsequently used to identify a minimum mean validation error and the associated training epoch was used to build the final model weights.

c. Predictive Model Comparisons

Comparative analyses were conducted to assess the model accuracies relative to benchmarks, including a “featureless baseline” [46], and the SoA prediction model. The equations used for implementing the SoA model, as well as the list of inputs to each model, including our linear model, can be found in Table 2 (expansion on the implementation of the Zhang et al. weighted time series prediction is available in supplementary material). The featureless baseline error was generated for each of the ten folds in cross-fold validation where the average of all labels from the training set was used to create a naïve prediction (not using any feature data). Prediction error below the featureless baseline indicates that a model is superior to the simple guess of the mean of the population.

Table 2. Variables included in each prediction

	Metabolic								
	Baseline	Trial	Height	Speed	Age	Weight	Population	Terrain	Time
Featureless									
Zhang et al.		✓							2 min
Linear	✓	✓	✓	✓	✓	✓	✓	✓	1-2 min

Equation	Reference
$\begin{bmatrix} \widehat{y_1} \\ \dot{\hat{E}} \end{bmatrix} = A^{+}y$	Zang et al. [22]

$\widehat{y_1}$ and $\dot{\hat{E}}$ were estimates of the initial metabolic response and steady state metabolic rate. A^{+} was the pseudo-inverse of A which was an iteratively constructed weight matrix based on the number of metabolic readings in the time window, and y was the vector of metabolic costs measured in that window.

In order to apply the weighted time series prediction model used by Zhang et al., we implemented the metabolic power estimation equation from the supplementary material of their previously published work [22]. However, we observed that the model consistently underestimated steady state metabolic power from our training data. To address this discrepancy and conduct the best comparison, several terms were recalibrated for our specific dataset. This involved five rounds of grid searching based on the provided constants [22], [47]. The width of the search range was progressively reduced in each round of tuning to allow the identification of more refined constants. We confirmed that the tuning process effectively minimized the error in unimpaired level-ground walking to comparable mean ranges (5.2 ± 2.0%) as previously reported (4.3%). Next, additional tuning was performed on all remaining training data (inclusive of participants with walking impairment).

d. Data Analysis

We established flexible criteria for the inclusion and processing of walking trials in our dataset used to train our linear models so that our models would be robust to experimental variation when applied prospectively. Each trial required a standing metabolic baseline period from which to calculate the net metabolic power of the walking trial. The metabolic baseline was calculated as the mean metabolic power recorded while quietly standing; during this time the participant was not allowed to speak or make any large physical movements. The duration of the quiet standing period varied depending on the study; 3 minutes was the most common duration. We then defined an active walking trial period that commenced at the start of each walk and progressed to include two or more minutes of steady state measurements. Walking trials were generally 6-8 minutes in duration, with steady state selected from the last two minutes. The shortest trial was five minutes and the longest was 16 minutes. In each trial, steady state was confirmed visually by two separate researchers. Slight variation in the selected start time of each walking trial was purposefully allowed to replicate how the model would be applied to a real-world testing scenario where the initial application of the model could vary slightly relative to the first step a participant takes. Time series measurements of expired volumetric oxygen consumption and carbon dioxide production were recorded by the metabolic system for each trial. We calculated metabolic power using the modified Brockway equation (equation 1)

$$\text{Metabolic Power} = (VO_2 * 16.5835 + VCO_2 * 4.51)/(60 * BM) \quad (1)$$

where VO_2 was the volume (mL) of oxygen consumption, VCO_2 was the volume (mL) of carbon dioxide production, and *BM* was each participant's body mass in kg [48].

Following the calculation of metabolic power from equation 1, the time series of metabolic power for each walking trial was interpolated to estimate values at every second within the trial duration. This interpolation was essential due to the inconsistent sampling rate of multiple devices and allowed for more accurate segmentation in the machine learning training data. The ranges observed for sampling rates in this study were between one sample every second to one sample every 23 seconds. The most common sampling frequency was one sample every three seconds, and the second most common was at ten second intervals. Interpolation on each raw time series was implemented using three different techniques, resulting in three interpolated time series that were then used in model training so as to increase the size of our dataset and introduce small variations (see supplementary material). The first method was a repeated observations approach, which filled timepoints between observations with the last observed value and did not alter the fidelity of the recorded data, keeping it functionally equivalent to raw data. The second and third methods linearly interpolated between each recorded timepoint to fill out the array. The third method differed from the second method by using the standard deviation of the steady state metabolic power to add simulated noise to every added point. By processing each input observation with all three interpolation methods, variations in the values assigned to ramp rate and labels were also introduced. This approach allowed us to effectively triple the size of the dataset by generating three distinct input trials from a single input trial, thus enhancing

the robustness of our models. Our ramp rate metric was calculated based on 120 samples after interpolating ~10 to ~100 recorded measurements (depending on the sampling frequency). Importantly, data in the test set were not linearly interpolated, so that testing was only done on measured data [49].

Our input data included participant characteristics and experimental condition variables. Participant characteristics included their population (CP, elderly status, stroke history, or unimpaired), height (centimeters), age (years) and weight (kilograms). Experimental condition variables included walking speed (meters/second), walking condition (overground, level treadmill, incline treadmill), and assistance type (none/shod, exoskeleton zero-torque, exoskeleton assistance, ankle-foot orthotics). In all cases where an exoskeleton was included (zero-torque or assistance) the exoskeleton was an untethered bilateral ankle exoskeleton weighing between 2.2 and 2.8 kg, and providing 0.1-0.35Nm/kg of peak assistive torque [12], [14], [30], [31], [32], [33].

3. Withheld Data Validation

To increase the generalizability of our results, a withheld data set was excluded from all model training/testing, and was never viewed by the algorithm until after final training was complete. The withheld data set was selected intentionally to include each of our participant populations and assistive device conditions (Table 3). Each withheld data set trial represented a unique participant, with one exception for the single stroke participant who had two trials included.

Table 3. Withheld test set information

	Shod	Zero-torque	Assistance	AFO
CP I	1	1†	1	0
CP II	0	0	1*, 1	0
CP III	1†	1*	1*	1
Elderly	1	1	0	0
Stroke	1	0	1	0
Unimpaired	0	1*	2	0

† overground trials
* incline trials
CP I-III refers to Gross Motor Function Classification Score Levels I, II, or III, respectively.

4. Results

Through the stepwise linear regression we found that the significant features were limited to the aggregate “Ramp Rate” ($p < 0.001$), height ($p < 0.001$), weight ($p < 0.001$), baseline ($p = 0.019$), walking speed ($p = 0.027$), CP I-III ($p < 0.001$), incline ($p < 0.001$), zero torque ($p = 0.009$), overground ($p < 0.001$), and the weighted time series prediction ($p < 0.001$).

We found that our linear model, using only significant features identified by the stepwise regression, predicted steady state metabolic power with 94.8±1.0% accuracy. This was within a single standard deviation of the observed mean measurement error and overlaps with the mean reported accuracy (95.7%). The prediction errors associated with the withheld dataset were not significantly different from the testing error ($p = 0.675$). The overall best SoA (weighted time series prediction) model was a significantly 4.1% ($p < 0.001$) more accurate than the one constrained by values used in the initial publications, additionally the linear model with significant features was a significant 3.3% ($p = 0.006$) more accurate than the optimally tuned weighted time series prediction (Fig. 2).

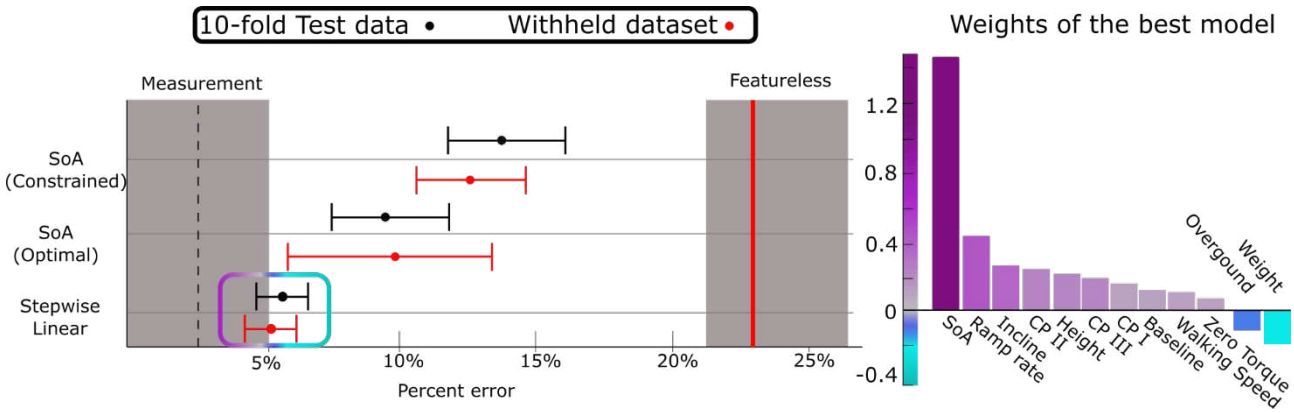

Figure 2. The featureless error for folds is shown with the grey background (22-27%), and the featureless accuracy on the withheld data is shown with a red line. The mean measurement error is shown with a dashed black line (2.49%) and the standard deviation is represented with a grey background on either side of the mean. The top row shows the accuracy of the SoA when constrained by published values, the middle row shows accuracy of the SoA when optimized for all training data, and the bottom row shows the accuracy of the linear model with significant features, each using the 2-minute time window. The best model is circled and the corresponding weights for all features are plotted to the right of the results. CP I-III represents Gross Motor Function Classificaiton Score Levels I-III respectively.

The accuracy of the best model at each time interval were examined on specific demographics of the test set (Fig. 3). Error for impaired populations in the withheld dataset improved to within the measurement error for all 5 of our examined populations

relative to the optimized weighted time series (SoA) prediction; CP I error decreased from 8.4% to 3.9%, CP II decreased from 14.4% to 2.5%, CP III decreased from 13.1% to 4.8%, Elderly error reduced from 14.5% to 3. 9%, and error in predictions for individuals with a history of stroke decreased from 9.3% to 1.2%.

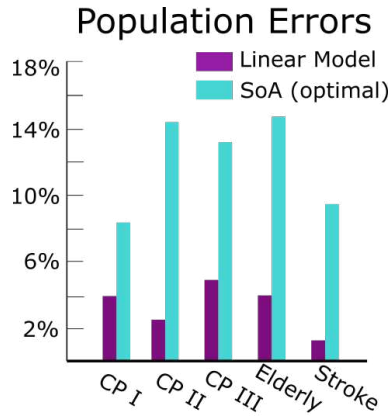

Figure 3. The linear model and the SoA were tested on the withheld data for clarity on which populations would be most able to use them. As expected, the SoA was not as accurate on unimpaired participants as compared to the accuracies reported in the literature. This was expected due to the modifications made to better predict on incline trials and in impaired populations. CP I-III represents Gross Motor Function Classificaiton Score Levels I-III respectively.

5. Discussion

Our primary goal was to validate an accessible predictive model for predicting steady state metabolic power during walking in individuals with gait deficits. The results of this study support our hypotheses that demographic and condition-specific information could be used to expand the prediction of metabolic power during steady state walking in individuals with neurological and walking deficits. Our second objective to validate the use of an aggregate metabolic measure (i.e., “ramp rate”) corresponded to the second

largest weighting factor in our final model. Finally, a comparison of our model with existing approaches in the literature demonstrated improved steady state metabolic power prediction accuracy for individuals with walking deficits. By training a model using data collected on individuals with movement impairments and while using assistive devices, we were able to produce a more accurate predictive equation for these patient populations and common walking conditions. The linear model consistently improved the accuracy of the state of the art, though the weighted time series prediction was the most heavily weighted feature.

The accuracies of the model proposed here demonstrated that it was able to predict steady state metabolic power with two minutes of data and contextual information about the trial with an accuracy of 94.8 ± 1.0% for all populations, and 95.3 ± 2.7% on populations with walking impairments. These results are in line with the level ground unimpaired metabolic prediction accuracies reported by Weyand et al. (which used weight, height, and walking speed) and Zhang et al. (which used metabolic data only). The Weyand model saw 91.9 ± 6.7% accuracy while Zhang saw an average accuracy of 95.7%, with a mode of 96.3%, and a minimum of 84.5% and was successfully able to optimize exoskeletons with human-in-the-loop techniques. Our sample size of 44 unique individuals is also within the range of these peer studies where Weyand et al. had 78 subjects and Zhang et al. had 7. This implies that our proposed method might be effective in providing predictions accurate enough for optimization studies to be conducted on impaired populations.

The linear model revealed a few interesting insights into the effects of, and interactions between, several model parameters. Some seemingly obvious effects were evidently captured by the model, like positive relationships between incline and speed on metabolic power. Our model also captured evidence of slightly increased cost of transport during treadmill vs overground walking [50]. Prior studies have demonstrated that the energy cost of walking increases with the severity of neurological deficits from CP (GMFCS levels I-III) [29]. Somewhat surprisingly, however, our model produced similar weighting factors for GMFCS levels I-III (labeled as CP 1-3), likely because other model inputs more accurately captured each individual's metabolic response (e.g., SoA, ramp rate, etc.). Similarly, we were surprised that exoskeleton assistance was not a statistically significant model input, likely for the same reason as was just mentioned. Features including age, unimpaired, stroke, AFO, shod, elderly, and stroke history were not identified as significantly impacting the accuracy of the linear model. The authors believe this most likely was due to another feature more accurately capturing the information presented by each of these variables (e.g., walking speed).

A core motivation was to develop and validate an easy to implement metabolic prediction model to facilitate assistive device development for individuals with neuromuscular conditions. As researchers in assistive device technology aim to reduce the metabolic burden of movement for patient populations like CP and stroke, they can utilize this model to more quickly evaluate the influence of device configurations on the energy cost of walking. The accuracy of our model should allow for the use of in-the-loop optimization of wearable assistive devices in individuals with limited walking

capacity [37]. To facilitate the utilization of our model, we have made available a standalone MATLAB executable that accepts the simple biometric inputs we have detailed here, and which can be paired with a real-time connection to a metabolic system for the streaming of metabolic rate data.

This study had several limitations that could be expanded upon in future work. First, our model used binary flags for assistance conditions, and the addition of contextual information, such as the amount of assistance or additional power being added to the user, may improve prediction accuracy. Another recommended future research investigation would be to explore whether crouch severity could improve prediction accuracy in CP. Also, trial data from individuals were allowed to be in both the training and testing sets. Ideally, we would have had enough participants in every category, shown in Table 1, to use participant separated data in each train and test set; however this was not feasible with the impaired populations examined at this time. Future validation work should test our model in real-time during prospective testing and evaluate model accuracy for participants absent in the training dataset. Next, more complex modeling paradigms could also be considered and layered to extract more complex interactions between parameters. Applying a convolutional deep neural network to the transient metabolic measurements could produce a second direct predictor of metabolic power at steady state, which could in turn be a feature in a linear model or subsequent deep neural network. The main reason this was not explored was the desire for interpretability of our model. Recently created models, such as Kolmogorov-Arnold Networks, could be used in future developments to maintain

interpretability and utilize machine learning's ability to identify relevant interactions whether linear or nonlinear [51]. Lastly, there were limitations in the data available for training. The dataset was noticeably missing unimpaired shod trials as well as having very little data on individuals with a history of stroke or AFO users.

In summary, this study builds on the weighted time series prediction of steady state metabolic power by demonstrating the benefit of modeling contextual information capturing a variety of experimental factors. Our interpretable model demonstrated promising results for predicting steady state metabolic power from two minutes of walking metabolic rate data individuals with CP and other patient populations. This work may unlock the use of in-the-loop optimization of wearable assistive devices in individuals with limited walking capacity.

6. Acknowledgments

We would like to thank all of the researchers who collected the original data for this study; Dr. Ying Fang, Dr. Greg Orekhov, and Dr. Jack R. Williams, as well as the non-author students who helped to filter and process individual trials from the hundreds of potential trials initially identified. This work was supported in part by the National Science Foundation under Grant 2045966. This work was also supported in part by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under award numbers R15HD099664, 1R01HD107277, and F30HD103318. The content is solely the responsibility of the authors and does not

necessarily represent the official views of the National Science Foundation or National Institutes of Health.

7. References

[1] D. B. Maltais, M. R. Pierrynowski, V. A. Galea, and O. Bar-Or, “Physical activity level is associated with the O2 cost of walking in cerebral palsy.,” *Med Sci Sports Exerc*, vol. 37, no. 3, pp. 347–53, Mar. 2005, doi: 10.1249/01.mss.0000155437.45937.82.

[2] K. M. Michael, J. K. Allen, and R. F. MacKo, “Reduced ambulatory activity after stroke: The role of balance, gait, and cardiovascular fitness,” *Arch Phys Med Rehabil*, vol. 86, no. 8, pp. 1552–1556, 2005, doi: 10.1016/j.apmr.2004.12.026.

[3] F. B. Ortega, J. R. Ruiz, M. J. Castillo, and M. Sjöström, “Physical fitness in childhood and adolescence: a powerful marker of health,” *Int J Obes (Lond)*, vol. 32, no. 1, pp. 1–11, Jan. 2008, doi: 10.1038/SJ.IJO.0803774.

[4] H. K. Graham *et al.*, “Cerebral palsy,” *Nat Rev Dis Primers*, vol. 2, no. 1, p. 15082, Dec. 2016, doi: 10.1038/nrdp.2015.82.

[5] J. R. Gage, M. H. Schwartz, S. E. Koop, and T. F. Novacheck, *The identification and treatment of gait problems in cerebral palsy*, 2nd ed. London: Mac Keith Press, 2009.

[6] K. M. Steele, A. Rozumalski, and M. H. Schwartz, “Muscle synergies and complexity of neuromuscular control during gait in cerebral palsy,” *Dev Med Child Neurol*, vol. 57, no. 12, pp. 1176–1182, 2015, doi: 10.1111/dmcn.12826.

[7] C. Lepage, L. Noreau, and P. M. Bernard, “Association between characteristics of locomotion and accomplishment of life habits in children with cerebral palsy,” *Phys Ther*, vol. 78, no. 5, pp. 458–469, 1998, doi: 10.1093/ptj/78.5.458.

[8] D. Maltais *et al.*, “Use of orthoses lowers the O2 cost of walking in children with spastic cerebral palsy,” *Med Sci Sports Exerc*, vol. 33, no. 2, pp. 320–325, Feb. 2001, doi: 10.1097/00005768-200102000-00023 LK - https://arizona-primo.hosted.exlibrisgroup.com/openurl/01UA/01UA?&sid=EMBASE&issn=01959131&id=doi:10.1097%2F00005768-200102000-00023&atitle=Use+of+orthoses+lowers+the+O2+cost+of+walking+in+children+with+spastic+cerebral+palsy&stitle=Med.+Sci.+Sports+Exerc.&title=Medicine+and+Science+in+Sports+and+Exercise&volume=33&issue=2&spage=320&epage=325&aulast=Maltais&aufirst=D.&auinit=D.&aufull=Maltais+D.&coden=MSCSB&isbn=&pages=320-325&date=2001&auinit1=D&auinitm=.

[9] B. C. Bennett, S. D. Russell, and M. F. Abel, “The effects of ankle foot orthoses on energy recovery and work during gait in children with cerebral palsy,” *Clinical Biomechanics*, vol. 27, no. 3, pp. 287–291, 2012, doi: 10.1016/j.clinbiomech.2011.09.005.

[10] B. B. *et al.*, “The effect of hinged ankle-foot orthosis on gait and energy expenditure in spastic hemiplegic cerebral palsy,” *Disabil Rehabil*, vol. 29, no. 2, pp. 139–144, Jan. 2007, doi: 10.1080/17483100600876740 LK - https://arizona-primo.hosted.exlibrisgroup.com/openurl/01UA/01UA?&sid=EMBASE&issn=03790797&id=doi:10.1080%2F17483100600876740&atitle=The+effect+of+hinged+ankle-

foot+orthosis+on+gait+and+energy+expenditure+in+spastic+hemiplegic+cerebral+palsy&stitle=Disabil.+Rehabil.&title=Disability+and+Rehabilitation&volume=29&issue=2&spage=139&epage=144&aulast=Balaban&aufirst=Birol&auinit=B.&aufull=Balaban+B.&coden=DREHE&isbn=&pages=139-144&date=2007&auinit1=B&auinitm=.

[11] G. Orekhov, Y. Fang, J. Luque, and Z. F. Lerner, “Ankle Exoskeleton Assistance Can Improve Over-Ground Walking Economy in Individuals with Cerebral Palsy,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 2, pp. 461–467, Feb. 2020, doi: 10.1109/TNSRE.2020.2965029 LK - https://arizona-primo.hosted.exlibrisgroup.com/openurl/01UA/01UA?&sid=EMBASE&issn=15580210&id=doi:10.1109%2FTNSRE.2020.2965029&atitle=Ankle+Exoskeleton+Assistance+Can+Improve+Over-Ground+Walking+Economy+in+Individuals+with+Cerebral+Palsy&stitle=IEEE+Trans.+Neural+Syst.+Rehabil.+Eng.&title=IEEE+Transactions+on+Neural+Systems+and+Rehabilitation+Engineering&volume=28&issue=2&spage=461&epage=467&aulast=Orekhov&aufirst=Greg&auinit=G.&aufull=Orekhov+G.&coden=ITNSB&isbn.

[12] Z. F. Lerner *et al.*, “An untethered ankle exoskeleton improves walking economy in a pilot study of individuals with cerebral palsy,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 10, pp. 1985–1993, 2018, doi: 10.1109/TNSRE.2018.2870756.

[13] Z. Gabriel and A. Bowling, “Quality of life from the perspectives of older people,” *Ageing Soc*, vol. 24, no. 5, pp. 675–691, 2004, doi: 10.1017/S0144686X03001582.

[14] Y. Fang, K. Harshe, J. R. Franz, and Z. F. Lerner, “Feasibility evaluation of a dual-mode ankle exoskeleton to assist and restore community ambulation in older adults,” *Wearable Technologies*, vol. 3, 2022, doi: 10.1017/wtc.2022.12.

[15] M. Raitor, S. Ruggles, S. L. Delp, C. K. Liu, and S. H. Collins, “Lower-limb exoskeletons appeal to both clinicians and older adults, especially for fall prevention and joint pain reduction,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 1577–1585, 2024, doi: 10.1109/TNSRE.2024.3381979.

[16] A. Lakmazaheri, S. Song, B. B. Vuong, B. Biskner, D. M. Kado, and S. H. Collins, “Optimizing exoskeleton assistance to improve walking speed and energy economy for older adults,” *J Neuroeng Rehabil*, vol. 21, no. 1, pp. 1–14, 2024, doi: 10.1186/s12984-023-01287-5.

[17] H. Mtaweh, L. Tuira, A. A. Floh, and C. S. Parshuram, “Indirect calorimetry: History, technology, and application,” *Front Pediatr*, vol. 6, no. September, pp. 1–8, 2018, doi: 10.3389/fped.2018.00257.

[18] P. Krustrup, A. M. Jones, D. P. Wilkerson, J. A. L. Calbet, and J. Bangsbo, “Muscular and pulmonary O2 uptake kinetics during moderate- and high-intensity sub-maximal knee-extensor exercise in humans,” *Journal of Physiology*, vol. 587, no. 8, pp. 1843–1856, 2009, doi: 10.1113/jphysiol.2008.166397.

[19] W. Felt, J. C. Selinger, J. M. Donelan, and C. D. Remy, "'Body-in-the-loop': Optimizing device parameters using measures of instantaneous energetic cost," *PLoS One*, vol. 10, no. 8, pp. 1–21, 2015, doi: 10.1371/journal.pone.0135342.

[20] P. Slade, M. J. Kochenderfer, S. L. Delp, and S. H. Collins, "Personalizing exoskeleton assistance while walking in the real world," *Nature*, vol. 610, no. 7931, pp. 277–282, 2022, doi: 10.1038/s41586-022-05191-1.

[21] J. R. Koller, D. H. Gates, D. P. Ferris, and C. D. Remy, "'Body-in-the-Loop'Optimization of Assistive Robotic Devices: A Validation Study.," in *Robotics: Science and Systems*, 2016.

[22] J. Zhang *et al.*, "Human-in-the-loop optimization of exoskeleton assistance during walking," *Science (1979)*, vol. 356, no. 6344, Jun. 2017, doi: 10.1126/science.aal5054.

[23] P. G. Weyand, B. R. Smith, N. S. Schultz, L. W. Ludlow, M. R. Puyau, and N. F. Butte, "Predicting metabolic rate across walking speed: One fit for all body sizes?," *J Appl Physiol*, vol. 115, no. 9, pp. 1332–1342, 2013, doi: 10.1152/japplphysiol.01333.2012.

[24] M. Kim *et al.*, "Human-in-the-loop Bayesian optimization of wearable device parameters," *PLoS One*, vol. 12, no. 9, pp. 6–8, 2017, doi: 10.1371/journal.pone.0184054.

[25] K. A. Ingraham, E. J. Rouse, and C. D. Remy, "Accelerating the estimation of metabolic cost using signal derivatives: Implications for optimization and evaluation of wearable robots," *IEEE Robot Autom Mag*, vol. 27, no. 1, pp. 32–42, 2020, doi: 10.1109/MRA.2019.2954108.

[26] S. Faraji, A. R. Wu, and A. J. Ijspeert, “A simple model of mechanical effects to estimate metabolic cost of human walking,” *Sci Rep*, vol. 8, no. 1, pp. 1–12, 2018, doi: 10.1038/s41598-018-29429-z.

[27] J. M. Argilés, N. Campos, J. M. Lopez-Pedrosa, R. Rueda, and L. Rodriguez-Mañas, “Skeletal Muscle Regulates Metabolism via Interorgan Crosstalk: Roles in Health and Disease,” *J Am Med Dir Assoc*, vol. 17, no. 9, pp. 789–796, 2016, doi: 10.1016/j.jamda.2016.04.019.

[28] P. K. Gill, K. M. Steele, J. M. Donelan, and M. H. Schwartz, “Causal modelling demonstrates metabolic power is largely affected by gait kinematics and motor control in children with cerebral palsy,” *PLoS One*, vol. 18, no. 5 May, pp. 1–15, 2023, doi: 10.1371/journal.pone.0285667.

[29] M. Nardon, F. Ruzzante, L. O’Donnell, A. Adami, S. Dayanidhi, and M. Bertucco, “Energetics of walking in individuals with cerebral palsy and typical development, across severity and age: A systematic review and meta-analysis,” *Gait Posture*, vol. 90, no. August, pp. 388–407, 2021, doi: 10.1016/j.gaitpost.2021.09.190.

[30] G. Orekhov, Y. Fang, J. Luque, and Z. F. Lerner, “Ankle Exoskeleton Assistance Can Improve Over-Ground Walking Economy in Individuals with Cerebral Palsy,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 2, pp. 461–467, 2020, doi: 10.1109/TNSRE.2020.2965029.

[31] Y. Fang and Z. F. Lerner, “Feasibility of Augmenting Ankle Exoskeleton Walking Performance with Step Length Biofeedback in Individuals with Cerebral Palsy,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 442–449, 2021, doi: 10.1109/TNSRE.2021.3055796.

[32] B. C. Conner and Z. F. Lerner, “Improving Ankle Muscle Recruitment via Plantar Pressure Biofeedback during Robot Resisted Gait Training in Cerebral Palsy,” *IEEE International Conference on Rehabilitation Robotics*, vol. 2022-July, pp. 25–29, 2022, doi: 10.1109/ICORR55369.2022.9896581.

[33] B. C. Conner, J. Luque, and Z. F. Lerner, “Adaptive Ankle Resistance from a Wearable Robotic Device to Improve Muscle Recruitment in Cerebral Palsy,” *Ann Biomed Eng*, vol. 48, no. 4, pp. 1309–1321, 2020, doi: 10.1007/s10439-020-02454-8.

[34] R. W. Armstrong, “Definition and classification of cerebral palsy,” *Dev Med Child Neurol*, vol. 49, no. 3, p. 166, 2007, doi: 10.1111/j.1469-8749.2007.00166.x.

[35] M. Noorkoiv *et al.*, “Predictors of Walking Efficiency in Children with Cerebral Palsy: Lower-Body Joint Angles, Moments, and Power,” *Phys Ther*, vol. 99, no. 6, pp. 711–720, 2019, doi: 10.1093/ptj/pzz041.

[36] J. M. A. Alcantara *et al.*, “Validity of four commercially available metabolic carts for assessing resting metabolic rate and respiratory exchange ratio in non-ventilated humans,” *Clinical Nutrition*, vol. 41, no. 3, pp. 746–754, 2022, doi: 10.1016/j.clnu.2022.01.031.

[37] J. Zhang *et al.*, “Human-in-the-loop optimization of exoskeleton assistance during walking,” *Science (1979)*, vol. 356, no. 6344, Jun. 2017, doi: 10.1126/science.aal5054.

[38] R. C. Chen, C. Dewi, S. W. Huang, and R. E. Caraka, “Selecting critical features for data classification based on machine learning methods,” *J Big Data*, vol. 7, no. 1, Dec. 2020, doi: 10.1186/s40537-020-00327-4.

[39] P. McCullagh, *Generalized Linear Models*. Routledge, 2019.

[40] T. M. H. Hope, “Linear regression,” *Machine Learning: Methods and Applications to Brain Disorders*, pp. 67–81, Jan. 2020, doi: 10.1016/B978-0-12-815739-8.00004-3.

[41] G. James, D. Witten, T. Hastie, R. Tibshirani, and J. Taylor, *Linear Regression: An Introduction to Statistical Learning.* . Cham: Springer International Publishing, 2023.

[42] J. F. C. M. C. X. S. C. H. H. W. E. Pan Zhou, “Towards theoretically understanding why sgd generalizes better than adam in deep learning.,” *Adv Neural Inf Process Syst*, vol. 33, pp. 21285–21296, 2020.

[43] I. K. Nti, O. Nyarko-Boateng, and J. Aning, “Performance of Machine Learning Algorithms with Different K Values in K-fold CrossValidation,” *International Journal of Information Technology and Computer Science*, vol. 13, no. 6, pp. 61–71, 2021, doi: 10.5815/ijitcs.2021.06.05.

[44] V. Pedregosa, F. and Varoquaux, G. and Gramfort, A. and Michel, P. and Thirion, B. and Grisel, O. and Blondel, M. and Prettenhofer, A. and and Weiss, R. and Dubourg, V. and Vanderplas, J. and Passos, and E. Cournapeau, D. and Brucher, M. and Perrot, M. and Duchesnay, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[45] P. Flach, “Performance Evaluation in Machine Learning: The Good, the Bad, the Ugly, and the Way Forward,” *33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial*

Intelligence, EAAI 2019, pp. 9808–9814, 2019, doi: 10.1609/aaai.v33i01.33019808.

[46] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, 2nd ed. Springer, 2009.

[47] J. C. Selinger and J. M. Donelan, “Estimating instantaneous energetic cost during non-steady-state gait,” *J Appl Physiol*, vol. 117, no. 11, pp. 1406–1415, 2014, doi: 10.1152/japplphysiol.00445.2014.

[48] J. M. Brockway, “Derivation of formulae used to calculate energy expenditure in man,” *Hum Nutr Clin Nutr*, vol. 41, no. 6, pp. 463–471, 1987.

[49] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.

[50] K. Parvataneni, L. Ploeg, S. J. Olney, and B. Brouwer, “Kinematic, kinetic and metabolic parameters of treadmill versus overground walking in healthy older adults,” *Clinical Biomechanics*, vol. 24, no. 1, pp. 95–100, 2009, doi: 10.1016/j.clinbiomech.2008.07.002.

[51] Z. Liu *et al.*, “KAN: Kolmogorov-Arnold Networks,” Apr. 2024, [Online]. Available: http://arxiv.org/abs/2404.19756