Optimizing the Parameters of A Physical Exercise Dose-Response Model

An Algorithmic Comparison

Mark Connor
Natural Computing Research And Applications Group,
University College Dublin
Dublin, Ireland
mark.connor@ucdconnect.ie

Michael O'Neill
Natural Computing Research And Applications Group,
University College Dublin
Dublin, Ireland
m.oneill@ucd.ie

ABSTRACT

The purpose of this research was to compare the robustness and performance of a local and global optimization algorithm applied to the problem of fitting the parameters of a non-linear dose-response model utilized in the field of exercise physiology. Traditionally the parameters of dose-response models utilised in exercise physiology have been fit with non-linear least squares procedures in combination with local optimization algorithms. These algorithms have demonstrated limitations in their ability to converge on a globally optimal solution. This research purposes the use of an evolutionary computation based algorithm as an alternative method to fit a non-linear dose-response model. The results of our comparison over 1000 experimental runs demonstrated the superior performance of the evolutionary computation based algorithm to consistently achieve a more consistent model fit and holdout evaluation performance in comparison to the local search algorithm. This initial research would suggest that global evolutionary computation based optimization algorithms are a fast and more robust alternative to local optimization algorithms when fitting the parameters of non-linear dose-response models.

CCS CONCEPTS

• **Applied computing** → Health care information systems;

KEYWORDS

Intelligent Systems In Sport, Athletic Training, Model Fitting, Differential Evolution

ACM Reference Format:
Mark Connor and Michael O'Neill. 2021. Optimizing the Parameters of A Physical Exercise Dose-Response Model: An Algorithmic Comparison. In *2021 Genetic and Evolutionary Computation Conference Companion (GECCO '21 Companion), July 10–14, 2021, Lille, France.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3449726.3459494

1 INTRODUCTION

It has been observed that many biological systems exhibit nonlinear characteristics by examining the relationship between an input to a biological system, typically known as a dose, and the subsequent system output or response. In the context of a dose-response model related to physical exercise, we consider the system inputs to be a combination of the volume and intensity of exercise activity, while the response is the cellular and tissue adaptions of the human body. A common example of this relationship is the cardiac tissue adaptations that occur in response to an increase in the volume of an aerobic exercise activity such as running, swimming or cycling.

2 DOSE-RESPONSE MODEL FITTING

We fit the following dose-response model which asserts that the response to a single exercise training dose has both a positive physiological response referred to as "fitness" and a negative response referred to as "fatigue" [2], represented mathematically in the following manner:

$$p_t = p_0 + K_1 \sum_{i=1}^{t-1} w_i e^{-\frac{n}{r1}} + K_2 \sum_{i=1}^{t-1} w_i e^{-\frac{n-i}{r^2}}, \quad n \in \mathbb{N}. \tag{1}$$

Where p_0 is an initial baseline level, $K1$ and $K2$ are the positive weighting factors for fitness and fatigue, these values represent the rate at which an individual can recover from a training stimulus, r^1 and r^2 represent the time decay until fitness and fatigue return to baseline, w_i is the exercise activity workload stimulus measured during a single session and n denotes the nth instant of time. To fit the parameters of the model and ensure that it returns robust and accurate predictions, we seek to minimise the differences between the predicted and observed responses using non-linear least squares regression (NLS) [4, 5, 7]. In this experiment, we compare the performance of a Limited Memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm [6] to a seeded and un-seed version of the Differential Evolution (DE) algorithm [9], to perform the minimization and model parameter optimization. To generate a seeded population of values for the DE Seeded option the initial randomised parameter values, which are provided to the L-BFGS algorithm, are used to generate a Gaussian distribution from which 19 random seeded values are selected in addition to the initial parameter values. The generated normal distribution had a mean value equal to its random initial parameter and a standard deviation equal to ± 5% of that parameter value. The purpose of generating the distribution in such a way was to provide the DE seeded algorithm with enough diversity to create new solution candidates to explore the search

---

GECCO '21 Companion, July 10–14, 2021, Lille, France

ACM ISBN 978-1-4503-8351-6/21/07.
https://doi.org/10.1145/3449726.3459494

space effectively but also to focus the area from which it started the exploration. The rationale for this procedure was to perform a closer comparison of the algorithms when the search process is initiated from a similar area of the search space and a randomly initialised area as is the case in the unseeded DE algorithm [1]. This process was repeated 1000 times using a new set of random initial parameter values during each experimental run. To test the fit of the optimized models an open-source data set consisting of 166 exercise training sessions and observed responses was used [3, 8].

3 RESULTS

The mean±SD model fit coefficient of determination (R^2) scores are as follows: DE ($M = 0.997, SD = 0.00$), DE-Seeded ($M = 0.976, SD = 0.070$), L-BFGS ($M = 0.877, SD = 0.201$). Figure 1 displays the loss scores (sum of the squared errors) between the predicted and observed response variables over the 1000 experimental runs using a hold out data. The results of a one-way analysis of variance (ANOVA) test comparing the fitted R-Squared scores are: $Fstatistic = 264.71$, $p < 0.001$, while the results of an ANOVA test using a hold out data set mean loss scores are: $Fstatistic = 90.07$, $p < 0.001$. These results would suggested that there is significant difference in the performance of the algorithms at an $\alpha = 0.05$.

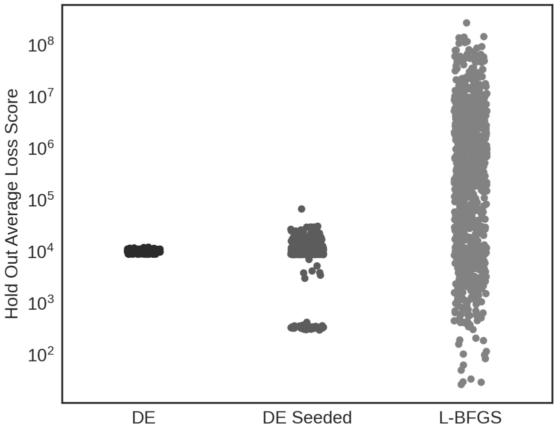

Figure 1: Sum of the squared errors between the fitted model predicted and observed response variables

Figure 2 displays the distribution of the optimization process run time for each algorithm during the experimental runs. The results demonstrate that the DE and DE seeded algorithm consistently achieve a strong model fit and low hold out set prediction error. The L-BFGS algorithm displayed higher variance and worse performance across the experimental runs.

4 CONCLUSIONS & FUTURE WORK

In this paper, we set out to compare the performance and robustness of a local versus global optimization algorithm when fitting the parameters of a non-linear dose-response model. Based on our observations we found that the differential evolution global algorithm is capable of repeatedly finding optimal model parameter values demonstrated by strong model fits and performances on a holdout data set. The local search L-BFGS algorithm displayed low robustness to varying initial parameter settings and had inferior performances when predicting response values on a holdout data set. A seeded version of the differential evolution algorithm displayed the best overall performance across all measures. These result would suggest that when fitting a dose-response model, of a similar type to that used in this research, a differential evolutionary algorithm seeded with reasonable initial parameters may provide a fast, robust and high performing alternative to local search algorithms such as L-BFGS. Future work should be conducted to confirm these findings using a wider set of algorithmic settings and variants, as well as larger more diverse real-world data sets.

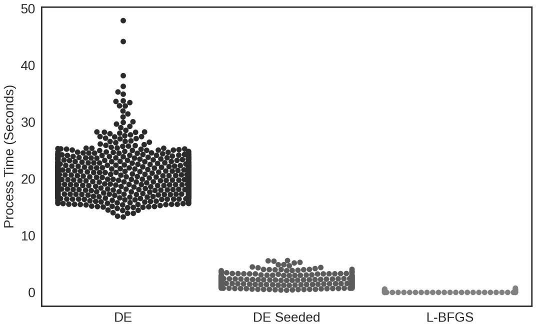

Figure 2: Process Running Time

REFERENCES

[1] Anthony Brabazon, Michael O'Neill, and Seán McGarraghy. 2015. *Natural Computing Algorithms.* Springer Berlin Heidelberg, Berlin, Heidelberg.
[2] Thomas W. Calvert, Eric W. Banister, Margaret V. Savage, and Tim Bach. 1976. A Systems Model of the Effects of Training on Physical Performance. *IEEE Transactions on Systems, Man, and Cybernetics* SMC-6, 2 (1976), 94–102.
[3] David C Clarke, Philip F Skiba, and D C Clarke. 2013. A Personal View Rationale and resources for teaching the mathematical modeling of athletic training and performance. *Adv Physiol Educ* 37 (2013), 134–152.
[4] Philippe Hellard, Marta Avalos, Lucien Lacoste, Frederic Barale, Jean-Claude Chatard, and Gregoire P. Millet. 2006. Assessing the limitations of the Banister model in monitoring training. *Journal of Sports Sciences* 24, 5 (2006), 509–520.
[5] Muzaffer Kapanoglu, Ilker Ozan Koc, and Senol Erdogmus. 2007. Genetic algorithms in parameter estimation for nonlinear regression models: An experimental approach. *Journal of Statistical Computation and Simulation* 77, 10 (2007), 851–867.
[6] Dong C. Liu and Jorge Nocedal. 1989. On the limited memory BFGS method for large scale optimization. *Mathematical Programming* 45, 1-3 (1989), 503–528.
[7] Jun Ma, Eric Bair, and Alison Motsinger-Reif. 2020. Nonlinear Dose–Response Modeling of High-Throughput Screening Data Using an Evolutionary Algorithm. *Dose-Response* 18, 2 (2020).
[8] H. Motulsky and A. Christopoulos. 2004. Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting.
[9] R Storn and Price K. 1997. Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization* 11, 4 (1997), 341–359.