

ベイズ推定と深層学習による

ロバストなフルマラソン完走タイム予測モデルの構築と SHAP 値による解釈

梶浦 雅之[†] 稲葉 龍一郎[†] 宮崎 陽輔[†] 平野 賢[†]
 Masayuki Kajiura Ryuichiro Inaba Yosuke Miyazaki Ken Hirano

1. はじめに

フルマラソンレースでは様々なペース戦略が存在し[1], 自身の実力に応じたペース設定が重要である. さらに, 自身の実力に基づいて練習時のペースを設定することで, 練習効率の向上が期待される[2]. 以上より, フルマラソンレースにおける完走タイムの予測は, レース中の適切なペース設定を通じたペース戦略への活用に加え, 練習効率の向上に貢献する.

近年, フルマラソンの完走タイム予測のための特徴量として Critical Speed (CS) が注目されている. CS は長時間持続可能な最高速度として定義される速度であり[3], 最大努力下での様々な走行距離における速度をもとに算出される. フルマラソンは長時間の走行を伴うため CS と完走タイムの関連が報告されている[3]. 一方で, CS の算出には複数の最大努力下での走行を行う必要があり, 実験室と異なり統制が困難な実環境下でこれらを取得することは容易ではない. また, レース中のペース設定へ CS を活用するためにはレース直前に計測することでより適切なレース中のペース設定が可能である一方で, レース前のランナーに対して最大努力下で複数回の走行を課すことはレース時のパフォーマンスへの影響が大きいと容易ではない.

本稿では, 非最大努力下での走行が含まれた走行時の速度データよりベイズ推定を用いて推測した CS と, 練習状況から完走タイムを予測する深層学習モデルを構築し, SHAP 値を用いてモデルを説明した. 非最大努力下や実環境下での走行に由来するノイズに対して, ベイズ推定を用いることでロバストに CS を推定し, フルマラソンの完走タイムを予測した (Fig 1). 本研究では, 実環境下で取得された最大/非最大努力下の走行データが混在する場合でも, ロバストなフルマラソンの完走タイム予測を行うことで, フルマラソンのレース支援に貢献する.

2. 関連研究

実環境下での CS 推定には適切なモデル選択が必要である. Smyth らはフルマラソン 16 週間前までの実環境下で取得された特定の距離における走行距離と走行時間のデータを用いて線形回帰により CS の推定を行った[3]. この手法では完走タイムと CS の関連が報告された一方で, 実際のマラソンレースでは CS と比較して平均 (±標準偏差) 84.8% ± 13.6% の誤差が確認された. また, Kranenburg らはトレッドミルでの走行から推定された CS は屋外の競技場で取得されたものと比較して過大評価することを指摘している [4]. さらに, Patoz らはデータ取得方法に応じて CS を推定するモデルの選択が必要であると指摘[5]しており, 実環境下で計測された走行データをもとに CS を推定

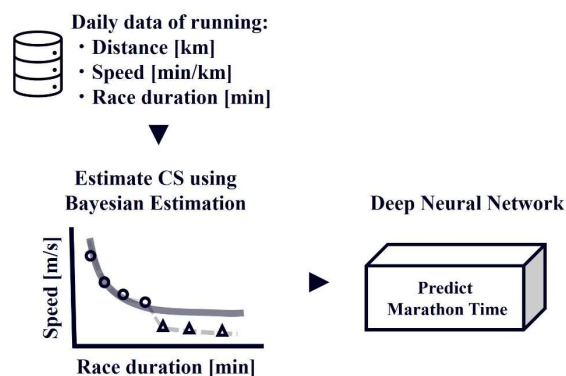


Fig 1. システムの概要

する場合は, 実環境下で発生しうるノイズを考慮した適切なモデルが必要であると考えられる.

一方で, CS 以外にも完走タイムと関連がある特徴量が報告されている. Yamaguchi らは月ごとの練習量や頻度と完走タイムとの関連を示した[6]. また, Pumares らは高レベルなランナーは低強度の走行を指す Zone1 の割合が全体の走行に対して大きくなる傾向を示した[7]. 以上より, 本研究では日常的に取得された走行データのノイズを考慮し CS を推定することに加えて, 練習の量や質を考慮した完走タイムの予測を行うことで非最大努力下の走行を用いた場合においても, 最大努力下の走行を用いて完走タイムを予測した場合と同様の精度での予測を目指す.

レースの完走タイム予測は様々な方法を用いて行われている. Till らはトレッドミルの傾斜を利用してフルマラソンの完走タイムの予測を行った[8]. また, Knechtle らは BMI や皮下脂肪の厚さを計測することで, ハーフマラソンレースの完走タイムの予測を行った [9]. さらに, Lerebourg らは 10km の最大努力下での走行記録と性別, BMI のデータを用いて Neural Network や K 近傍法などの機械学習の手法によりフルマラソン完走タイムの予測を行った[10]. これらの手法は高い予測精度を示す一方で, 特殊な機器による測定やレース直前の最大努力下の走行を必要とするため, より簡便な予測手法が求められる. また機械学習を用いた完走タイムの予測において, 予測タイムに対する特徴量の寄与を説明することはコーチング現場など, より応用的な分野において活用が期待できる. そこで本研究では, 特殊なテストや機器による測定を必要とせずに完走タイム予測を行う深層学習モデルを構築することに加えて, SHAP 値を用いたモデルの解釈を行うことで, より応用的な活用を目指す.

[†] Institute of sport science, ASICS corporation

3. データ

2023年3月から2025年1月にかけて312（男性：287，女性：24，その他：1）人の練習時の走行データ（平均：132件，標準偏差：73件）を取得した．走行データはGPSを内蔵し最小250mから1kmごとに計測可能なウェアラブルデバイス（CMT-20S-AR，カシオ計算機株式会社）を通じて一定距離ごとの走行速度（ラップデータ）や距離，時間が取得された．また，Webサービスを通じて各ランナーの過去2年間におけるフルマラソンの最高記録を取得し，これを完走タイム予測の正解データとした．自己最高記録は平均230.2，標準偏差34.2 [min]であった．また取得した自己最高記録に基づいて，各ランナーのレベルをSub3，Sub3.5，Sub4，Sub4.5，Sub5の5段階に割り当てた．Sub3は自己最高記録が3時間未満であることを示す．Sub3，Sub3.5，Sub4，Sub4.5，Sub5，Sub5.5はそれぞれ21，72，120，63，21，15人であった．

4. 解析

4.1 前処理

各ランナーの練習状況として，月間で最大の走行距離や平均走行速度，また以降に導出するCSとD'に加えて，CSを用いて算出される低強度と高強度の練習の割合（Zone1，Zone3）を説明変数とした．Zone1，Zone3は先行研究[7]を参考に，推定されたCSに対して各走行データの平均速度が82.3%以下，100%以上の走行回数の割合とした．加えてCSの推定のために，各走行データに対して，1km~10kmまで1kmずつにおける最速の走行速度を取得した．各最速の走行速度は移動平均法を用いて，各ラップデータに対して窓のサイズを1km~10kmまで1kmずつ変化させることで，各距離における平均の走行速度を取得し，窓のサイズに応じてランナーごとに最速の走行速度を取得した（Fig 2）．

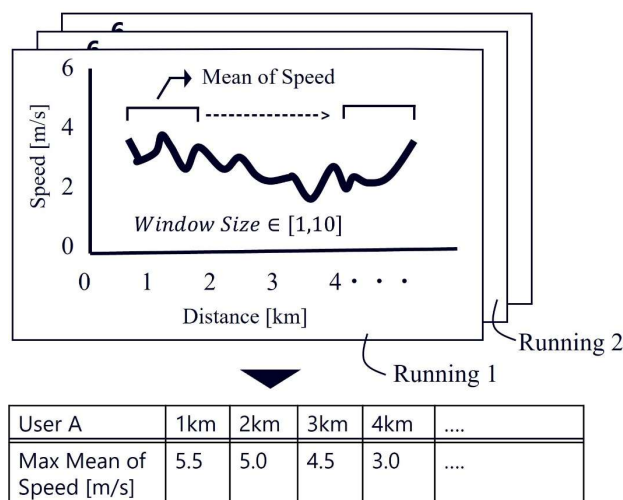


Fig 2. 最速の走行速度取得の流れ

4.2 ベイズ推定を用いたCSの導出

CSは複数距離に対して最大努力で走行した際の速度を用いて推定が可能である一方で，本研究で利用するデータは日常的な練習時の走行速度を用いるため，最大努力では

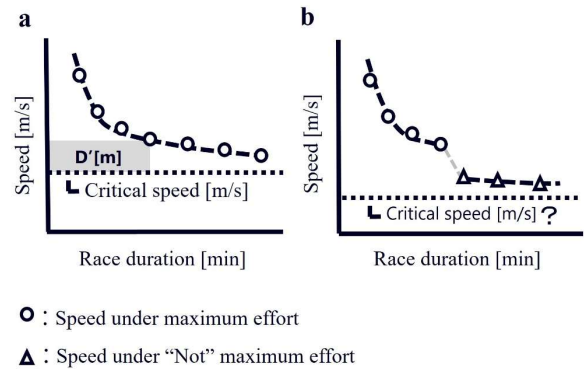


Fig 3. 走行時間(Race duration)に対する走行速度 (Speed) : a) 最大努力での走行のみ ; b) 非最大努力での走行が含まれる場合

ない走行時のデータが含まれる．Fig. 3a, 3bに最大努力/非最大努力での走行データが含まれる場合の走行時間に対しての速度の変化の例を示す．Fig. 3aでは走行時間に対して，走速度がなめらかに単調減少している様子が予想される一方で，Fig. 3bでは走行時間が長くなると急激に走速度が減少する場合が発生する．以上の傾向は，練習における長時間の走行では最大努力での走行ではなく各ランナーが快適に走行可能な速度で走行するというランナーの練習様態によって発生すると考えられる．理想的な最大努力下での走行時間と走行可能な速度の関係式は以下の二点が条件として考えられる：

- ・走行時間の増加に対して速度が単調減少
- ・漸近線が取得可能

以上より，本稿では指数関数を用いて，各ランナーより取得した走行時間に対する速度の近似式を算出し，漸近線をCSとして取得した．推定に使用した関数は以下である：

$$\text{Speed} = \exp(a \cdot \text{Raceduration}) + b \dots \text{式 1}$$

Race durationは走行時間 [min] で Speedは速度 [m/s]，aは減少の度合い，bは漸近線の値を示す．a,bの事前分布は一樣分布においてステップ幅を限定した分布を用いた．また，観測誤差は標準偏差0.5の半正規分布に従うと仮定した：

$$\varepsilon \sim \text{HalfNormal}(0,0.5)$$

以上より，尤度関数と a,b の同時事後確率を以下のように定めた：

$$\mathcal{L}(\text{Race duration, Speed} | a, b, \varepsilon) = \prod_i N(\text{Speed}_i | b + \exp(a \cdot \text{Race duration}_i), \varepsilon) \dots \text{式 2}$$

$$\begin{aligned} &P(a, b | \text{Race duration, Speed}) \\ &\propto \mathcal{L}(\text{Race duration, Speed} | a, b, \varepsilon) \\ &* \text{Uniform}(b | 1, 8) * \text{Uniform}(a | -0.3, -1e4) \dots \text{式 3} \end{aligned}$$

$\mathcal{L}$ は尤度を表し， P は同時事後確率を示す． $\text{Uniform}(a)$ と $\text{Uniform}(b)$ は事前分布，**Race duration**と**speed**は前処

理を経て得られた観測値を表し、添え字 i は 1km から 10km までの距離を示す。パラメータの推定は MCMC (Markov Chain Monte Carlo) 法の一つである NUTS (No-U-turn sampler) により a, b の事後分布を取得した。本研究では、5000 回のイテレーションを行い、バーンイン区間として 3000 回分のデータを除いたトレースプロットより収束を判断した。フルマラソンの完走タイムの特微量として、事後分布として取得された a, b の平均値を用いて近似式を導出することで CS と CS を超えて走行可能な距離として定義される D' (Fig 3a) を取得した。収束の判定は R -hat が 1.1 以下であることを確認した。また、CS 推定の評価のため、CS とフルマラソンの完走タイムを用いてピアソンの積率相関係数 (r) と最小二乗法を用いた線形回帰を行い、決定係数 (R^2) より CS の推定を評価した。

4.3 モデル構築と SHAP 値を用いた解釈

前処理を経て得られた特徴量に対して正規化を行い、フルマラソンの完走タイム予測を行うモデルを構築した。モデルは中間層が 3 層でニューロン数が 512 個、活性化関数を ReLU 関数とした全結合ニューラルネットワーク (深層学習) を用いた。層の数やニューロンの数、最適化関数に関しては Oputuna[11]を用いて決定された。またデータの偏りが起こらないようにするため、ランナーのレベルを均等に抽出する層化 K 分割交差検証を行った。データセットはそれぞれ重複のないように訓練用 80%, 評価データとして 20% に 5 度分割することで層化 K 分割交差検証を行った。評価指標として MAE (平均絶対誤差) と MAPE (平均絶対誤差率) を算出した。

構築された深層学習モデルでは説明変数に対してフルマラソンの完走タイム予測が出力される。一方で、それぞれの説明変数がどのように予測タイムの出力に寄与しているかを可視化することでそれぞれの説明変数の重要度を考察する。本研究では、SHAP (Shapley Additive exPlanations) [12]を用いて各予測タイムに対する説明変数の寄与度を算出した。また、各説明変数の完走タイム予測に対しての SHAP 値をレベルごとに可視化することで、モデルに対する大局的な解釈とレベルごとでの局所的な解釈を行った。解析は Python (ver. 3.12.8) と PyMC (ver. 5.20.0), SHAP (ver. 0.47.1) を用いて行われた。

5. 結果と考察

5.1 ベイズ推定を用いた CS の取得

Fig 4 にベイズ推定を用いて得られた CS とフルマラソンの完走タイムの関係及び、最小二乗法を用いた線形回帰による近似直線を示す。ピアソンの積率相関係数 (r) と決定係数 (R^2) はそれぞれ -0.79, 0.63 であった。フルマラソンの完走タイムに対して CS は負の相関を示し、特微量として活用可能であることが考えられる。Fig 5 に、あるランナーにおける観測値と推定された曲線を示す。観測値では Race duration が 25 分前後より急激な走行速度の減少が観られる一方で、推定された曲線はなめらかな単調減少が確認できる。本研究で用いた走行速度や距離などはスマートフォンの GPS データからも同様に取得が可能であることから、提案した手法は特殊な機器を用いず完走タイム予測が可能であると考えられる。

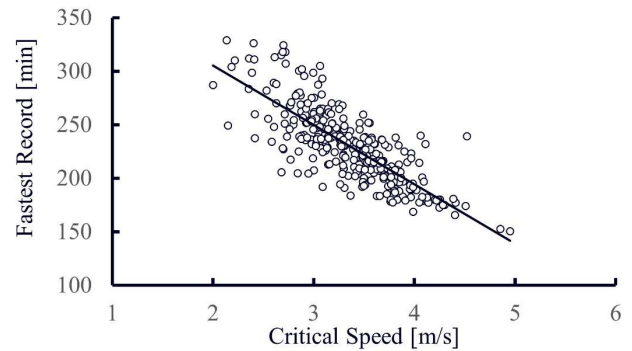


Fig 4. 推定された CS とランナーの完走タイムの関係

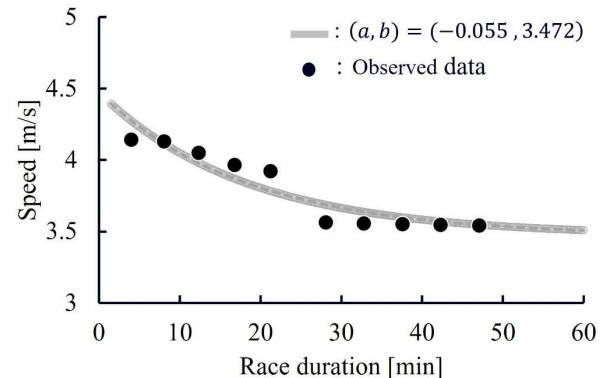


Fig 5. 観測値と推定された近似式の結果例 (図内の a, b はそれぞれ式 1 でのパラメータを示す)

5.2 完走タイム予測結果と SHAP 値を用いたモデルの解釈

予測モデルを通じて得られた MAE と MAPE はそれぞれ 13.5 [min], 5.9% であった。先行研究においてもフルマラソンの完走タイム予測は絶対誤差率が 10% 程度であり [8], 本手法では非最大努力が想定される走行データを用いて同程度の完走タイム予測が可能であることを示した。以上の結果は、提案した手法を用いて様々なレベルに対しても同一の枠組みで完走タイムの予測が可能であることを示唆している。今後の精度向上に向けては、各レベルのランナーは様々な練習様態を持つため、学習させるランナーの練習状況の多様性を考慮することが必要であると考えられる。実用化に向けては、データ数を増加させ、多様なランナーの練習状況を継続的に学習することで、様々な練習パターンに対する予測精度の向上が期待される。

Fig 6 に構築したモデルに対して SHAP 値を算出した結果を示す。横軸は SHAP 値を示しており、負の方向へ大きくなるほど完走タイムを速く予測することに寄与している。またプロット内の濃淡は値の大きさを示しており、色が濃くなるにつれて値が小さいことを示す。図より、CS は完走タイムに対して寄与が大きく、CS が大きいほど完走タイムが速くなると予測する傾向が見られた。また、月間の最大走行距離などの距離のデータは完走タイム予測への寄与が小さかった。Fig 7 に各レベルのそれぞれの特徴量に対する SHAP 値を示す。CS や平均走行ペースはレベルに応じて、予測タイムとの関連が観察された。また、

Zone1 の割合は、高レベルのランナーに SHAP 値が小さく完走タイムを速く予測する傾向が観られた。以上は、今回構築したモデルにおいて、練習の構成割合がフルマラソンの完走タイム予測に寄与していると考えられる。先行研究内でも練習構成と完走タイムの相関が報告されている[6]一方で、機械学習モデル内において練習構成が完走タイム予測に寄与することを示した研究は少ない。本研究の結果は、走行ペースや走行距離、走行速度に加えて練習構成を考慮することで完走タイム予測の精度向上が可能であり、さらにランナーのレベルによって完走タイムに寄与する項目が異なることを示唆する。

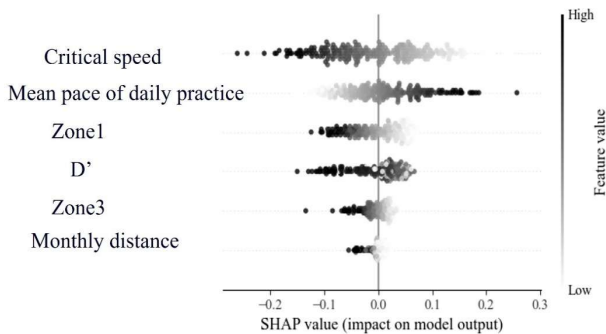


Fig 6. 構築した深層学習モデルに対する SHAP 値

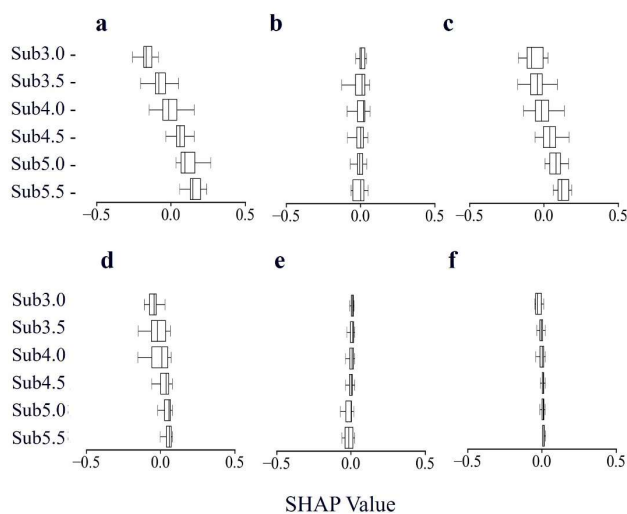


Fig 7. レベルごとの SHAP 値の箱ひげ図（各箱ひげ図において中央の箱は四分位範囲(IQR)を示し、右端のひげは第三四分位数 $+1.5 \times IQR$ 、左端のひげは第一四分位数 $-1.5 \times IQR$ を示す）：a) CS；b)D'；c) Mean pace of daily practice；d) Zone1；e) Zone3；f) Monthly distance

6. おわりに

本稿では、最大/非最大努力が混在する走行データに対して、ベイズ推定を用いた CS の推定を行った。また、フルマラソンの完走タイム予測に関する深層学習モデルを構築し SHAP 値を用いて各特徴量の寄与度を観察することでモデルの解釈を試みた。結果、予測誤差が 5.9%程度で完走タイムの予測が可能であり、Zone1 で示すような練習構成の特徴量が完走タイム予測に関連していることが示唆された。今後は、フルマラソン以外の距離のレースにおける完

走タイムを予測するモデルを構築する。また、完走タイム予測の値と SHAP 値による予測の説明を分かりやすく提示することで、より実用的なレース支援システムの構築を目指す。

参考文献

- [1] Abbiss, C.R., Laursen, P.B., Describing and Understanding Pacing Strategies during Athletic Competition, Sports Medicine., Vol.38 (2008).
- [2] Esteve-Lanao J, Foster C, Seiler S, Lucia A., Impact of training intensity distribution on performance in endurance athletes, Journal of Strength and Conditioning Research, vol.21(2007)
- [3] Smyth B, Muniz-Pumares D., Calculation of Critical Speed from Raw Training Data in Recreational Marathon Runners, Medicine & Science in Sports & Exercise, vol .52 (2020)
- [4] Kranenburg KJ, Smith DJ., Comparison of critical speed determined from track running and treadmill tests in elite runners, Medicine & Science in Sports & Exercise, vol .28 (1996)
- [5] Patoz A, Spicher R, Pedrani N et al., Critical speed estimated by statistically appropriate fitting procedures, European Journal of Applied Physiology, vol .121 (2021)
- [6] Yamaguchi A, Shouji M, Akizuki A et al., Interactions between monthly training volume, frequency and running distance per workout on marathon time, European Journal of Applied Physiology, vol.123 (2022)
- [7] Muniz-Pumares D, Hunter B, Meyler S et al., The Training Intensity Distribution of Marathon Runners Across Performance Levels, Sports Medicine, vol.55 (2025)
- [8] Till ES, Armstrong SA, Harris G et al., Predicting Marathon Time Using Exhaustive Graded Exercise Test in Marathon Runners, Journal of Strength and Conditioning Research, vol.30 (2016)
- [9] Knechtle B, Barandun U, Knechtle P et al., Prediction of half-marathon race time in recreational female and male runners, Springerplus., 16;3:248 (2014)
- [10] Lerebourg L, Saboul D, Cléménçon M et al., Prediction of Marathon Performance using Artificial Intelligence, International Journal of Sports Medicine, vol.44 (2023)
- [11] Takuya Akiba, Shotaro Sano, Toshihiko Yanase et al., Optuna: A Next-generation Hyperparameter Optimization Framework, KDD '19: The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Proceedings (2019)
- [12] Lundberg, Scott M. and Lee, Su-In, A unified approach to interpreting model predictions, Proceedings of the 31st International Conference on Neural Information Processing Systems (2017)