

## XGBoost とランダムフォレストによる高低狂いの変位予測 Displacement prediction of disorder of surface with randomforest and XGBoost

山田 泰亮<sup>†</sup> 森 康久仁<sup>‡</sup> 須鎗 弘樹<sup>‡</sup> 川本 一彦<sup>‡</sup>  
Taisuke Yamada Yasukuni Mori Hiroki Suyari Kazuhiko Kawamoto

### 1. はじめに

列車が線路の上を繰り返して走行する際の荷重によって、軌道の各部に変形や狂いが生じる。その変位量を数値化したものを軌道狂いと呼ぶ。軌道狂いが進行すると、列車の動揺や乗り心地に影響を与え、さらに進行の具合が大きくなると脱線の原因になることもある。

東海道新幹線ではドクターイエローと呼ばれる軌道検測者を約 10 日周期で走行させ、軌道狂いを数値として測定している。軌道狂いが進んだ線路の保線工事は、安全性を保つために定められた整備基準値と測定値を目視で比較することで、計画が立てられている。整備基準値に達していない箇所でも、急な軌道狂いの進行が起きる可能性があるため、次の測定を待たずに工事を行うこともある。軌道狂いの急な進行を急進と呼ぶ。軌道狂いの進行には多くの要因が関連し、進行の速度は測定地点や、様々な条件によって異なっている。そのため、軌道狂いの進行性は単純な線形回帰の外挿だけでは予測が難しいだけでなく、急進が起きることにより、整備基準値と実際の測定値のみでの判断が困難である。ドクターイエローの測定値から、次の測定値を予測することが出来れば、計画的な工事を行うことが出来るようになり、安全性の向上につながる。また、工事の平準化も期待される。

本研究では、軌道狂いのひとつである高低狂いを予測する。高低狂いとはレール登頂部の長さ方向の凹凸のことをいう。今回実験に用いるドクターイエローの測定データには、高低狂いを含む軌道狂い、ドクターイエローの測定時の上下動揺、速度など 10 種類のデータが含まれている(以後ドクターイエローデータと呼ぶ。)。また、工事が実施された箇所や日付が記録されたデータも用いる(以後、工事データと呼ぶ。)

三和ら[1]は、長い区間(100m)毎での高低狂いの時系列予測を行っている。本研究では、決定機を用いた機械学習である XGBoost[2]とランダムフォレスト[3]を用いて、ドクターイエローデータから次の測定時の 1m 範囲の短い区間での高低狂いの値を予測する方法を提案する。実験では、実際の測定値と予測された値を比較することで評価している。

## 2. XGBoost とランダムフォレストによる予測

XGBoost とランダムフォレストを用いて、ドクターイエローデータと工事データから次回測定値を予測する。

### 2.1 実験に使用するデータ

実験には、表 1, 表 2 に示された項目が記録されているドクターイエローデータと工事データを使用する。ドクターイエローデータに記録されている軌道狂いは、長さ 10m の弦をレールにあて、その中央における減とレールの離れを測定する 10m 弦正矢法で測定されている(図 1)。ドク

ターイエローデータは 340 回の測定分のデータが、東京駅から下り方向の 27000m~42000m の範囲で、1m 間隔で記録

されている(図 2)。工事データには、表 2 にある工事が行われた日付とその範囲が記録されている。工事の内容によって、影響を与える軌道狂いの種類や工事が行われる範囲が異なる。高低狂いは基本的に、マイナス方向に進行していき、工事が行われることで狂いを戻すように大きくプラスに変位する。

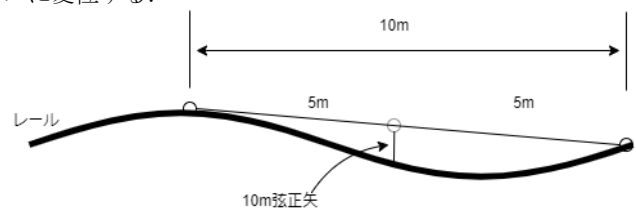


図 1 10m 弦正矢法

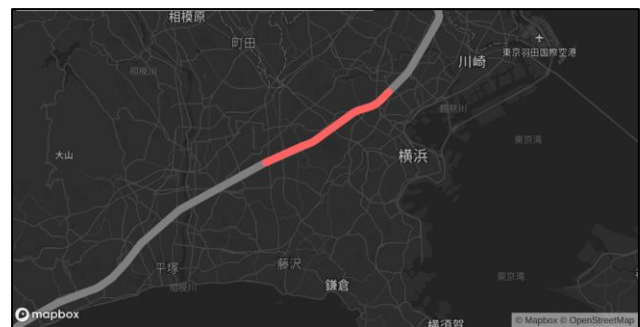


図 2 使用する測定区間  
(©Mapbox, ©OpenStreetMap)

#### 2.1.1 実験条件

予測の際の目的変数を以下の式に示す高低狂いの差分とし、予測に用いる特徴量としてドクターイエローデータの他の測定項目、1~4 回前の目的変数を用いる。

訓練データを 2011 年 4 月 6 日~2017 年 4 月 26 日までのデータ、テストデータを 2017 年 5 月 8 日~2020 年 9 月 25 日までのデータとして訓練データとテストデータを分割する。また、データ量を増やすために、訓練データには、同じ日付に測定された前後 1m の測定箇所のデータも加えている。

XGBoost とランダムフォレストのハイパーパラメータは optuna[4]を用いて調整する。Optuna の探索回数は 100 回とし、XGBoost とランダムフォレストに対して入力したハイパーパラメータとその探索範囲を表 3 に示す。XGBoost においてはハイパーパラメータのひとつである booster で dart と gbtree のどちらが選ばれるかによって使用するハイパーパラメータの種類や数が異なる。

本実験では、予測の対象として高低狂いの進行が整備基準値に近づいている測定箇所から 5 か所(東京駅から下り方向 27196m,27430m,29167m,30386m,31622m 地点)を選択してその結果を確認する。

また工事によって大きく変位した箇所を学習しないようにするために、何かしらの工事が行われた次の測定時刻のデータは削除し、学習には使用しない。

表 2 ドクターイエローデータ

データ属性	単位	データ属性	単位
最後尾の上下動揺	[g]	高低狂い(右レール)	[mm]
最後尾の左右動揺	[g]	水準狂い	[mm]
平面性狂い	[mm]	通り狂い(左レール)	[mm]
軌間狂い	[mm]	通り狂い(右レール)	[mm]
高低狂い(左レール)	[mm]	測定時の速度	[Km/h]

表 2 工事データに含まれている工事項目

むら直し	まくらぎ交換
バラスト作業	分岐器更換
レール更換	絶縁継目更換
簡易噴泥処理	伸縮継目更換

### 2.1.2 実験結果と考察

実験の結果を以下に示す(表 3)。各キロ程の結果に対して、MAE, MSE, RMSE の 3 つを出力する。

表 3 それぞれの測定箇所の実験結果

測定箇所	使用したモデル	MAE	MSE	RMSE
27196m	XGBoost	0.31	0.19	0.43
	ランダムフォレスト	2.09	5.63	2.37
27430m	XGBoost	0.07	0.01	0.09
	ランダムフォレスト	2.49	9.35	3.06
29167m	XGBoost	0.18	0.07	0.26
	ランダムフォレスト	2.76	9.45	3.07
30386m	XGBoost	0.25	0.12	0.35
	ランダムフォレスト	1.5	4.56	2.13
31622m	XGBoost	0.27	0.14	0.38
	ランダムフォレスト	1.95	4.91	2.22

それぞれの評価指標でランダムフォレストより、XGBoost の方がよい値を示していることが分かる。評価指標の値はいいが、この手法の問題として、予測対象の高低狂いの差分がとても小さい値であることが挙げられる。予測対象が小さい値であるため、0 に近い値を出力するだけで誤差が小さくなることも考えられる。

また、ランダムフォレストと XGBoost のそれぞれにラグモデルの傾向がある箇所が、何か所か見ることが出来るため正確な予測をしていると言い切ることはできない。

### 3. おわりに

本実験では、XGBoost とランダムフォレストを用いた高低狂いの変位を予測するシステムを提案している。本論文の手法では、ドクターイエローデータと工事データから、約 10 日後の次の測定日での値を予測することで、局所的な工事の計画を立てることにつながる。本実験の結果からは、XGBoost とランダムフォレストの二つを比較した際には、XGBoost の方が予測に有用である結果が得られている。

今後の研究展開として、使用するデータの追加や、線路の構造物の違いによる予測モデルの変更が挙げられる。使用するデータに関しては、降水量など、地面や軌道狂いに影響を与える現象のデータを追加することが考えられる。本研究では、トンネルや切取など線路がどの構造物の上に敷かれているのかには注目せずすべての測定箇所と同じ予測方法を使用している。しかし、線路の構造物にあったモデルを構造物ごとに変更することでその測定箇所にあった予測ができるようになると思われる。また、10m 弦正矢法で測定されている観点から見ると、測定箇所前後 1m のデータだけでなく、前後 5m 地点のデータを使用するなどデータの追加なども研究の展開内容として挙げられる。

提案手法によって、高低狂いの変位を予測することはできたが、実用化するには各測定箇所それぞれで高い精度が求められる。上にあげた研究展開などをおこない更なる精度の向上を目指す。

#### 謝辞

本研究を進めるにあたり、実験を行うための必要なデータを提供して下さった東海旅客鉄道株式会社に心から御礼申し上げます。

#### 参考文献

- [1] 三和雅史, 石川達也, 大山達雄, 軌道状態推移予測モデルの構築と最適軌道保守計画のための全整数型数理計画モデル分析, 土木学会論文集 No.681/IV-52, pp.51-65, 2001
- [2] Tianqi Chen, Carlos Guestrin. XGBoost: A Scalable Tree Boosting System (2016)
- [3] Breiman, L.: Random Forests, Machine Learning, Vol.45, No.1, pp.5-32 (2001).
- [4] Akiba, Takuya, et al." Optuna: A Next-generation Hyperparameter Optimization Framework." Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 2019.

† 千葉大学大学院融合理工学府 Graduate School of Science and Engineering, Chiba University

‡ 千葉大学大学院工学研究院 Graduate School of Engineering, Chiba University