

通り変位の時系列変化予測手法の検討

○ [土] 松田 博之 [土] 瀧川 光伸 (東日本旅客鉄道株式会社)

Study on Prediction Method of Time-Dependent Change of Alignment

○ Hiroyuki Matsuda, Mitsunobu Takikawa, (East Japan Railway Company)

By predicting track irregularity in the future using the track irregularities in the past measured with a track measurement device installed on trains in operation, this will get more precise and more efficient maintenance. In this research, the past track irregularity waveforms were utilized, the degradation characteristic for every place has been understood, and future track irregularity waveforms were predicted. Concretely, the time-dependent change prediction was performed from alignment using a regression model and support vector machine as measured by the track inspection car done in the previous year.

キーワード：通り変位，経時変化，予測，回帰モデル，サポートベクターマシン

Key Words：Alignment, Time-Dependent Change, Prediction, Regression Model, Support Vector Machine

1. はじめに

JR 東日本では，営業列車に検測装置を搭載して軌道変位を高頻度に検測する技術開発を進めている．高頻度検測による状態監視に加えて，蓄積した過去のデータを活用して将来の軌道変位を予測することで，より確実に効率的なメンテナンスの実現が期待できる．軌道変位進みの予測については，これまで軌道変位指数 P 値や標準偏差などの区間統計量を用いた将来予測手法が検討されてきた¹⁾．これに対して，本研究では，過去の軌道変位波形データを活用して場所毎の劣化特性を把握し，将来の軌道変位を波形のまま予測する手法を検討した．具体的には過去1年間の軌道検測車による通り変位データから単純回帰モデルおよび多次元回帰モデルを用いて時系列予測を行い，精度を検証した．

2. 対象データ

対象データは，軌道検測車 East-i により測定した 10m 弦通り変位データ (12 km 分) である．3 ヶ月毎に測定した 5 回分のデータを使用し，そのうち 4 回分を予測用データ，最後の 1 回を検証用データとした．検討区間はバラスト区間 (バラストと構造物との境界部を含む) であり，検測対象期間内に保守作業が行われた区間は含まれていない．

軌道検測車の検測データ (サンプリング間隔 0.25m) は，検測毎に位置ずれが生じるため，変位の予測を行う際に，変位測定位置の位置ずれ補正を行う必要がある．位置ずれ補正に関しては図 1 に示すように線形補完で補正した．

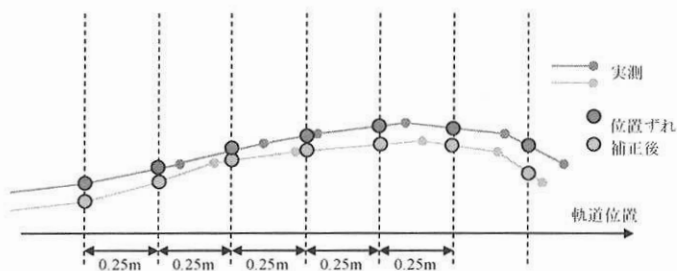


図1 位置ずれ補正の方法

3. 単純回帰モデル予測

3.1 予測手法

単純回帰モデル予測手法の概要を図 2 に示す．時刻 t ，任意の位置 i において検測された値を x_i^t とする．位置 i ごとに時刻 $x_i^{t-1} \sim x_i^{t+k-1}$ を既値として用いて， x_i^{t+k} を予測する．本検討では $k=5$ とし， $k=1 \sim 4$ を用いて x_i^{t+5} を予測した．

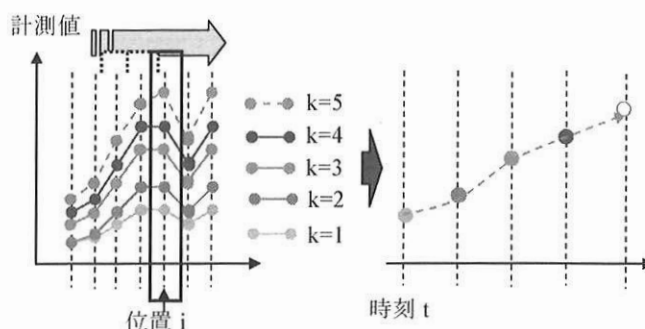


図2 単純回帰モデルの予測手法の概要図

3.2 予測結果

(1) 予測誤差の評価

本検討では4時点目までのデータをもとに5時点目の変位を予測する。予測値と5時点目の実測値との誤差を以下の2つの視点から評価する。

- 誤差の絶対値・・・予測値と実測値の誤差が何 mm 以内に収まったかという指標。
- 過去の変位に対する誤差の絶対値・・・誤差の絶対値が 1mm であったとしても、その大きさの評価は、その地点における過去の変位量に依存する。すなわち、過去に全く変位がなかった地点に対して誤差 1mm の予測をするのと、過去に大きく変位が起きている地点に対して誤差 1mm の予測をするのでは予測精度の価値が異なる。そこで、「変位差の平均値」に対する予測誤差の評価を行った。「変位差の平均値」は次のように定義される。

$$D = \frac{\sum_{i=1}^4 |d_{i+1} - d_i|}{4}$$

d_i : i 時点での変位

予測誤差を変位差の平均値で割った「変位差の平均値に対する予測誤差」は次のように定義される。

$$e = \frac{|y_{\text{predict}} - y_{\text{real}}|}{D}$$

y_{predict} : 5 時点目の変位の予測値

y_{real} : 5 時点目の変位の実測値

各地点について D に対する e の値を計算し、D の階級ごとの e の平均値を算出し、予測誤差の評価指標とした。

(2) 予測結果

図3に各条件下における単純回帰予測を行った場合の予測誤差指標 e を示す。変化の小さい区間、大きい区間にかかわらず予測誤差指標は 1.0 を上回っており（平均移動量の倍以上）、軌道変位の予測手法として単純回帰による予測精度はよくないといえる。

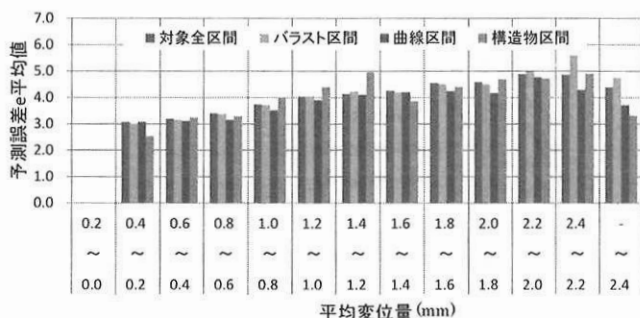


図3 単純回帰モデル予測結果

4. 多次元自己回帰モデル予測

4.1 予測手法

多次元自己回帰モデルによる予測は、1)モデル設定と2)回帰による予測式構築の2段階からなる。以下に、各段階に関して概説する。

(1) モデル設定

多次元自己回帰モデルは単純回帰モデルを拡張したものであり、軌道上のある1点のみだけでなく、隣接する複数の点の情報をもとに将来のある1点を予測する(図4)。本モデルでは、2つのパラメータを設定する必要がある。1つは「自己回帰の次数」すなわち過去のデータをどのくらいまで遡ってモデルに取り込むかというステップ数のパラメータである。もう1つは「空間方向の刻み幅」すなわち隣接する何点までモデルに取り込むかというパラメータである。これらは、少なすぎるとモデルに取り込む情報が少なくなり、多すぎてもノイズになってしまうため、適切なモデル設定が重要である。

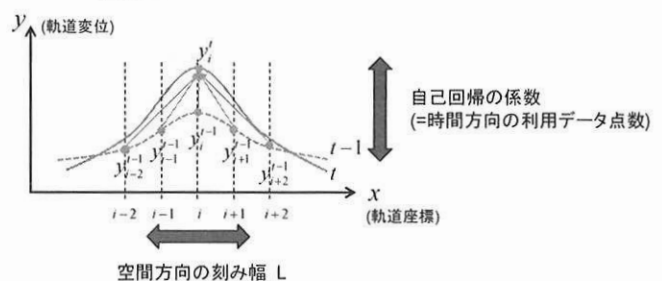


図4 多次元自己回帰モデル

(2) 回帰による予測式構築

本検討では、回帰分析のアルゴリズムとして Sparse Bayesian Learning (SBL) を用いた。SBL の特徴は、カーネル関数を用いた「非線形化」と、特徴的な点のみを用いて予測式を構築する「スパース化」である。

非線形化とは、データを一度非線形関数に入力して、その出力を利用して回帰を行うことである。線形回帰を行うとデータから大きくずれた直線の回帰式が得られてしまう。一方、データを一度非線形関数に入力し、その出力を用いて非線形な回帰を行うと、データに対して柔軟に合わせようとするため、精度よく回帰を行うことが可能になる。

次に、スパース化について説明する。単純な非線形回帰では予測式を構築する際に、全てのデータを用いて予測式を定式化している。一方、SBL を用いると、非線形回帰を行う際に予測式を特徴づける特定のデータのみを用いて予測式を構築する（全てのデータを回帰に用いるが、結果の式に用いるデータを選別する）。予測を行う際にこの特定のデータ点とそれに関連する重みのみを用いて計算を行うため、計算量が大幅に削減できる（スパース化）。

4.2 予測結果

(1) 予測誤差の評価

3.2 節 (1) 項同様、各地点について D に対する e の値を計算し、D の階級ごとの e の平均値を算出した。

(2) モデルのパラメータ設定

4.1 節 (1) 項で述べた 2 つの係数を変更し予測誤差を比較した。比較検討を行った区間は 1 km、係数は表 1 の 4 パターンである。表 1 のパラメータで多次元自己回帰を行い、図 5 の予測誤差の絶対値のヒストグラムを作成した。

パターン 1~3 に関しては、予測誤差のヒストグラムにほとんど差がない。空間方向のデータ点数を増やしても予測精度に大きな差がないため、空間方向には少なくとも両隣のデータを保持していれば十分であることがわかった。一方、パターン 4 は他のパラメータと比較して予測誤差が大きくなった。これは、今回の時系列データは時間方向の間隔が 3 ヶ月毎と大きいため、時間方向の情報を増やしても時間ごとの変化の規則性が薄れ、かえってノイズになってしまうためと考えられる。ただし、より短い間隔で時系列データが取得できれば、時間方向の変化の規則性が生きてくる可能性があるため、時間間隔が短い時系列データを用いる場合には検討の余地がある。以上の検討を踏まえ、今回はパターン 1 のパラメータを採用した。

表 1 モデルのパラメータのパターン

	自己回帰の次数	空間方向の刻み幅
パターン 1	1 (2 点分)	1 (3 点分)
パターン 2	1 (2 点分)	2 (5 点分)
パターン 3	1 (2 点分)	3 (7 点分)
パターン 4	2 (3 点分)	1 (3 点分)

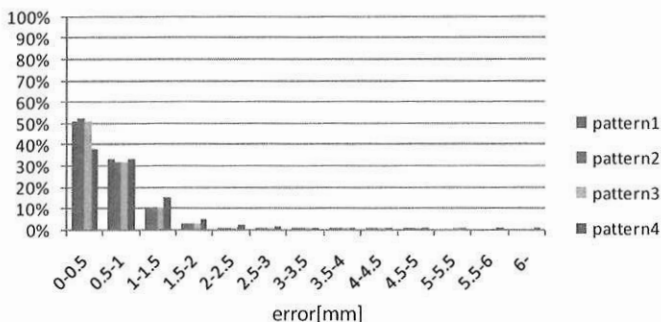


図 5 パラメータのパターンと予測誤差の絶対値の比較

(3) 予測結果

図 6 に各条件下における多次元回帰予測を行った場合の予測誤差指標 e を示す。全体的に、単純回帰手法よりも予測精度が改善した。一方、曲線区間、バラスト区間において $D=2.0$ 以上の予測精度低下が確認された。

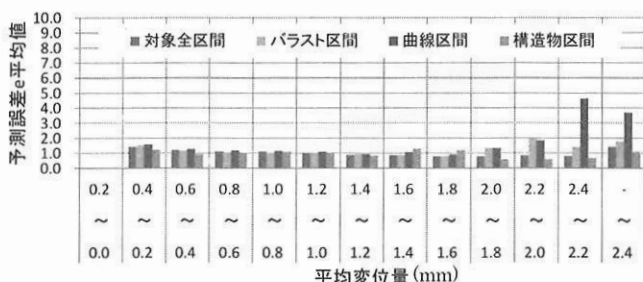


図 6 多次元回帰モデル予測結果

解析対象区間 (12 km) における誤差の絶対値のヒストグラムを図 7 に示す。予測誤差が 1 mm 以内の場所が 80% を超えており、ある程度の精度で予測ができていたことがわかった。一方、全体における割合は 1% 未満ではあるが、予測誤差が 6 mm 以上を超える箇所も数箇所あった。これらの箇所を確認すると、何れも過去の変位差の平均値が大きい箇所、すなわち急激に変化している箇所であることがわかった。

以上より、予測誤差が 1 mm 以内の箇所が 80% を超えたため、「多次元自己回帰モデル+SBL」は変位の漸増区間に関しては有効であることがわかった。一方、過去の変位差の平均値 D が大きいところ、すなわち変位の急進区間に関しては十分に予測精度を発揮できないことがわかった。変位が急進するかどうかを判定するためには別のアプローチが必要であると考えられる。

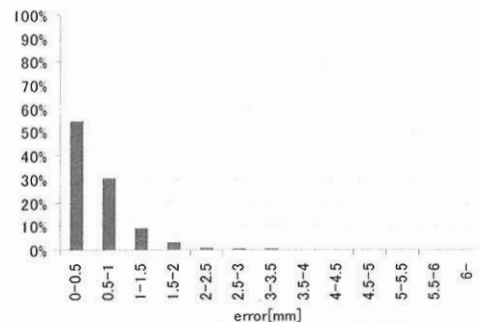


図 7 誤差の絶対値のヒストグラム

5. 変位急進区間に対する考察

変位急進区間を検知するために、非線形判別手法として近年着目されている Support Vector Machine (SVM) を用いた。本手法では、変位急進が起きた箇所の過去のデータを用いて、別の箇所も変位急進が起きるか起きないかを判別する。

5.1 変位急進の定義

変位が急進したという判定を行うために、本検討では図 8 のようにある地点に関して 4 時点目までの変位の変化量の平均値に比べて、4 時点目から 5 時点目までの変位の変化量が極端に大きい箇所を「急進箇所」と定義した。具体的には i 時点目から $(i+1)$ 時点目までの変位の絶対値を $d_{(i+1)(i)}$ としたときに

$$\text{急進率} = \frac{d_{54}}{(d_{43} + d_{32} + d_{21})/3} > 300\%$$

であるような箇所を急進箇所と定義した。

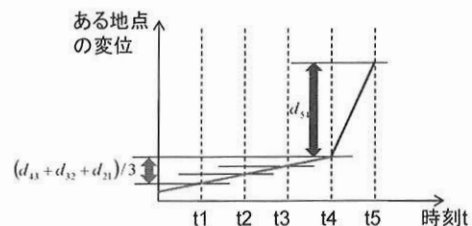


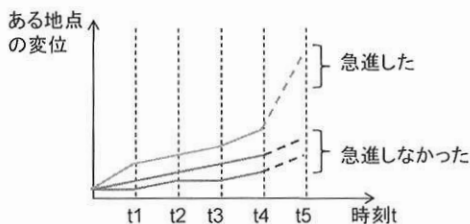
図 8 急進区間の定義

5.2 SVM による急進区間の判別手法

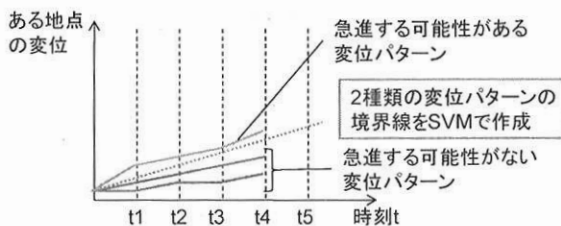
急進を確認するデータとして A 区間 (1km) と B 区間 (1km) を選定した。A 区間は急進箇所が全区間の 7% (=31 箇所) あり、急進率が 1100% を超える箇所があるため、急進検知が必要である。B 区間は急進箇所が全区間の 36% であり、急進率が 1000% を超える箇所がある。

今回は B 区間での急進結果をもとに判別基準を生成し (B 区間の急進データと似ているところを急進とする)、A 区間の変位の急進を検知する。急進区間検知の手順を図 9 に示す。なお、この判別に関しても多次元自己回帰モデル同様に、時間方向と空間方向にデータ点をとった (表 2)。

- ① まず B 区間の 5 時点目までのデータを使って、急進判定を行い、4 時点目までの変位パターンについてラベル (急進した・しなかった) をつける。



- ② 次に 4 時点目までの変位データと①でつけたラベルをもとに、急進する変位パターンと急進しない変位パターンの境界線を SVM で作成する。



- ③ 最後に、B 区間のデータから得られた境界線をもとに、A 区間の 4 時点目までのデータが今後急進するか判別する。

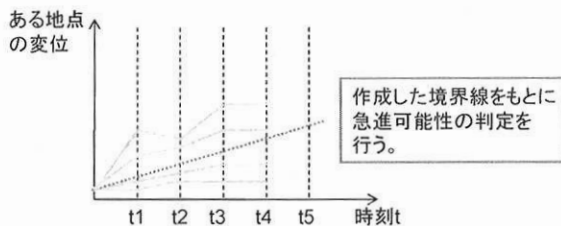


図 9 急進区間の検知アルゴリズム

表 2 データ点の取り方のパターン

	時間方向	空間方向
パターン I	3 点分	3 点分
パターン II	3 点分	5 点分
パターン III	3 点分	7 点分
パターン IV	2 点分	3 点分

5.3 変位急進区間の判別結果

SVM と B 区間のデータを用いて、A 区間において今後変位が急進すると予測した地点と、実際に変位が急進した地点との正当率を確認した (表 3)。今回は表 2 の各パター

ンに対して「急進検知率」と「誤警報率」で評価を行った。

結果 (表 4)、パラメータのパターン II、III において、正当率が 100% となり、B 区間のデータを用いて A 区間の急進を高い精度で検知できることがわかった。これは B 区間は急進箇所が全体の 36% であり、急進するバリエーションが A 区間よりも多かったため、A 区間の急進パターンを全て網羅できたと考えられる。

パターン I と II・III を比較すると、空間方向にデータ点を増やすことで急進検知率が上がるが、同時に誤警報率も上がることもわかった。また、時間方向のデータが少ないパターン IV では正当率が他のパターンと比べて極端に低くなった。これより急進する特徴を得るためには、時間方向のデータ点数をある程度確保する必要があると考えられる。

表 3 判別の正誤

	急進しないと判定	急進すると判定
実際に急進しなかった	未急進検知	誤警報
実際に急進した	見落とし	急進検知

表 4 急進判別結果

	パターン I	パターン II	パターン III	パターン IV
急進検知率	97%	100%	100%	3%
誤警報率	25%	30%	30%	22%

(※A 区間において、急進した: 31 箇所、しなかった: 369 箇所)

6. まとめ

- (1) 通り変位を予測する上では、単純回帰モデルよりも隣接する点の情報が利用できる多次元回帰モデルの方が予測精度が高く適した手法である。
- (2) 多次元自己回帰モデルを用いる際には、空間的、時間的な条件設定が予測精度に影響を与えるため、十分な検討が必要である。今回はパラメータを数パターン試してから検討を行ったが、実用化にあたってはこのパラメータを自動的に決定する枠組みが必要と考える。
- (3) 多次元自己回帰モデルは、変位が漸増している区間の予測には有効であるが、変位の急進区間に関しては十分に予測精度を発揮できない。急進区間を検知するには、多次元自己回帰モデルとは別のアプローチが必要と考える。
- (4) 変位急進区間の判別手法として SVM を用いた結果、高い精度で急進箇所を検知できた。今後も引き続き実データによる検証を行うとともに、誤警報率の低下方法や急進の判別基準生成のためのデータの効率のよい取り方について検討していく予定である。

最後に、本検討にご協力いただいた関係各位に感謝の意を表す。

参 考 文 献

- 1) 川口昭人, 三和雅史, 寺田浩一郎: 通り狂い進みの実態の把握と予測モデルの構築, 鉄道総研報告, Vol.19, No.2, pp.41-46, 2005.2