

## センシングデータを活用した車両台車部への着雪量推定モデルの開発

鉄道本部  
技術企画部  
**沢田 悠**  
現：鉄道本部 車両部



鉄道本部  
技術企画部  
**松田 篤史**



金沢支社  
新幹線車両課  
**遠藤 康一郎**



### 1. はじめに

車両台車部への着雪量が増加すると落下する可能性があるため、雪落とし作業が必要となります(図1)。現状は、気象情報会社から配信される降雪予報に基づいて雪落とし作業者の発動要否を判断していますが、発動した際に現車の着雪状況を確認すると、想定より着雪が少なく結果的に雪落とし不要となる状況が散見されています。

そこで本件では、気象条件や走行条件等のデータから車両台車部への着雪量を直接推定するより高精度なモデルを作成し、雪落とし作業要否判断支援ツールを導入することで、雪落とし作業の発動頻度最適化および着雪状況確認に伴う列車遅延の最小化を目指しています。



図1：北陸新幹線車両への着雪の様子

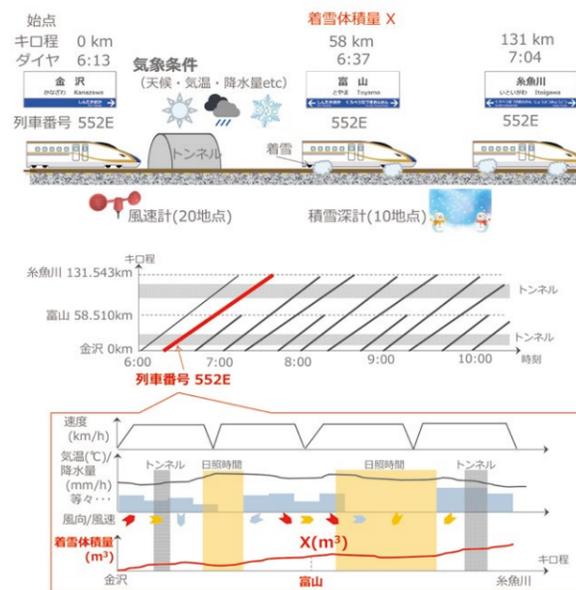


図2：着雪量推定モデル作成の条件設定

### 2. 概要・効果

#### (1) 着雪量推定の概要

金沢から東京方面に向かう北陸新幹線上り列車を対象とし、当日午前4時時点で得られる気象データから富山駅到着場面の各列車の着雪量を推定しました。着雪量推定モデル作成には、株式会社 SIGNATE と連携してオープンイノベーション形式のコンペティションを開催し、推定精度の高い上位3件を採用することとしました。(コンペティションサイト：<https://signate.jp/competitions/58>)

#### (2) 着雪量推定モデル作成用データセット

着雪量推定モデルの作成には下記のデータを使用することとしました。(参考：図2)

##### ①目的変数

各日の列車番号ごとに、富山駅到着時にレーザスキャナを用いて測定した車両台車部の着雪量実績値

##### ②説明変数

###### 【気象データ】

気象庁で開示している金沢・富山・糸魚川における1時間ごとの気象実績データ

###### 【列車データ】

列車番号毎の基本運転時刻および白山総合車両所発の列車番号情報

###### 【風速計・積雪深計データ】

沿線の1秒ごとの風速計、10分ごとの積雪深計の測定データ

###### 【地理データ】

土木構造物(停車場・トンネル)や風速計・積雪深計のキロ程、緯度経度情報

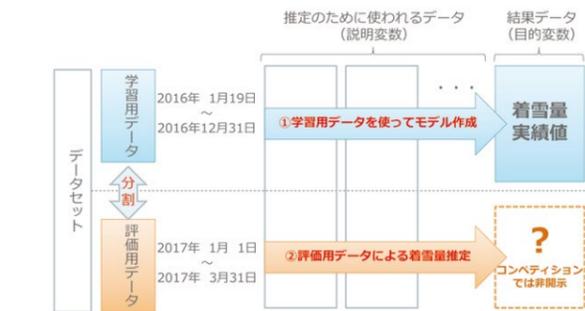


図3：学習用データと評価用データ

#### ③学習用データと評価用データ

本件では、説明変数と目的変数を提供して着雪量推定モデルを作成する学習用データと、作成した着雪量推定モデルの着雪量推定値を算出するために説明変数のみを提供する評価用データを準備しました。学習用データの提供期間は2016年1月19日～3月31日とし、評価用データの提供期間は2017年1月1日～3月31日としました(図3)。なお、評価用データの目的変数である着雪量実績値は、次項で述べる精度評価の際に使用しました。

#### (3) 推定精度評価方法

着雪量推定モデルの精度を定量的に評価するため、本件では、算出した列車番号毎の着雪量推定値と現車で測定した着雪量実績値の誤差の累積値を算出することとしました。

推定精度の評価関数にはWMAE(重み付き平均絶対誤差)を使用しました。着雪量の多い列車の推定精度をより重点的に評価するため、着雪量実績値が大きいほど重み付け係数が大きくなり、誤差の影響が大きくなる精度指標としました。推定精度が高いほどWMAEは小さな値となります。

$$WMAE = \sum_{i=1}^n w_i |y_i - \hat{y}_i|$$

$$w_i = \frac{1}{\sum_{j=1}^n (1 + |10^4 * y_j|)}$$

$y_{ij}$  : 着雪量実績値、 $\hat{y}_{ij}$  : 着雪量推定値(モデル)  
 $w_i$  : 重み付け係数

#### (4) 着雪量推定モデルの概要

推定精度の高かった上位3件の分析手法の概要は次の通りです。(図4)

##### ①1位モデル

- 学習用データから、着雪あり・なしのデータが同数となるようランダムに抽出し、3000パターンのミニ学習データを作成
- 3000パターンのミニ学習データに対して、3種類の着雪量推定モデル(Random Forest、XGBoost、AdaBoost)を作成(合計9000個のモデル作成)
- iiで作成した各推定モデルに対して評価用データの説明変数を入力することで、9000個の推定値が算出され、それらの中央値を着雪量推定値として出力

##### ②2位モデル

- 学習用データから、着雪あり・なしに二値分類するモデル(ロジスティック回帰)を作成
- 着雪ありの学習用データに絞って着雪量推定値を回帰分析(Random Forest Regressor)で算出するモデルを作成
- 評価用データの説明変数をi. およびii. のモデルに入力し、二値分類の推定値(確率)と回帰分析の推定値を出力。ここで、二値分類の推定値が予め定めた閾値以下となる場合は、回帰の推定値に関わらず推定値は0として出力

##### ③3位モデル

- 学習用データから、着雪量を回帰分析(XGBoost)で推定するモデルを作成
- 学習用データから、着雪量をディープラーニングによる判別分析(DNN Classifier)で推定するモデルを作成

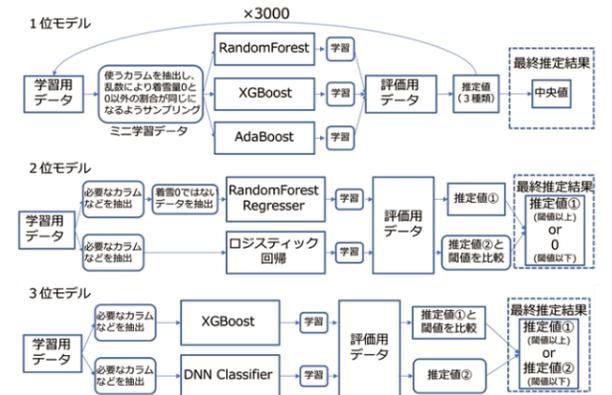


図4：着雪量推定モデルのロジック

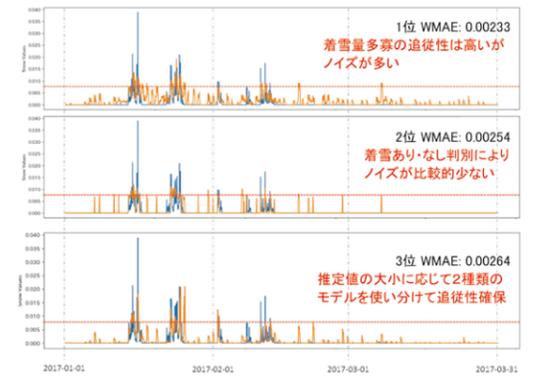


図5：着雪量推定モデル作成結果

【縦軸：着雪量、横軸：時系列】  
(橙線：推定値、青線：実績値、赤破線：雪落とし要否判断目安値)

- 評価用データの説明変数をi. のモデルに入力して着雪量推定値を出力し、予め定めた閾値と比較。閾値以上の場合はi. のモデル推定値を、閾値以下の場合にはii. のモデル推定値を出力

#### (5) 着雪量推定結果

上記3件の着雪量推定モデルを用いた着雪量推定値と着雪量実績値を比較検証しました(図5)。各モデルの分析手法による着雪量推定値の傾向の違いは見られますが、いずれのモデルにおいても着雪量実績値に対する一定の追従性を確保できたことから、午前4時時点で得られる気象条件等を説明変数に用いて、車両台車部の着雪量を直接推定するモデルの適用可能性を示唆する結果が得られました。

### 3. おわりに

今後は、実際の雪落とし作業要否判断場面で得られる気象予報と走行時の着雪量実績データを用いて毎冬再学習を重ね、さらなる推定精度向上と雪落とし作業要否判断支援ツールとしての実装化の取組みを継続していきます。