

イングランド一般道の AI による 短時間速度予測

宮崎 利行¹・緒方 陸²・村野 祐太郎³・菊池 恵和⁴・菅原 宏明⁴

¹正会員 八千代エンジニアリング株式会社 技術開発研究所
(〒111-8648 東京都台東区浅草橋 5-20-8 CS タワー)

E-mail: ts-miyazaki@yachiyo-eng.co.jp (Corresponding Author)

²非会員 八千代エンジニアリング株式会社 技術開発研究所

³非会員 八千代エンジニアリング株式会社 事業統括本部 国内事業部

⁴正会員 八千代エンジニアリング株式会社 技術開発研究所

交通状況のデジタルツインの活用方法として、短時間速度予測の検討を行った。イングランドの交通データはリアルタイムで公開されており、その過去データはオープンデータとしてダウンロード可能である。本研究ではダウンロードしたデータの中から一般道で比較的混雑している区間を選び、前後数地点との交通状況データの比較、機械学習 (Artificial Intelligence, 以下 AI) による速度予測を行った。合わせて SHapley Additive exPlanations (SHAP) による特徴量の重要度比較を行い、現況だけでなく過去の速度履歴にも予測性能改善の効果があることを明らかにした。また著者らの過去の検討では予測地点のデータのみを入力変数として短時間予測を行ったが、本研究では前後地点の交通状況を入力し、予測性能の改善が見られるか検討した。その結果、本研究の対象点では前後地点の交通状況を入力変数に加えても予測性能の改善は見られなかった。

Key Words: open data, traffic forecasting, machine learning, Shapley additive explanations

1. 概要

イギリスでは政府保有の National Highways 社がリアルタイムでイングランドの交通状況をリアルタイムで公開している¹⁾。その過去データは WebTRIS Traffic Flow API²⁾ からダウンロードして利用することができ、著者らはその活用法、特に短時間交通予測の検討を進めている³⁾。交通の「デジタルツイン」とでも言えるリアルタイムデータは日本でも国土交通省の xROAD (クロスロード)⁴⁾ での提供が検討されていたり、各社がプローブ交通情報として提供している。このようなデータは情報量が従来より格段に大きくなっており、特に複数地点の情報を組み合わせることでこれまで得られなかった情報を引き出せることが期待される。

短時間交通予測は Vlahogianni ら⁵⁾が現在と過去の交通情報に基づく数秒から 2, 3 時間先の予測と定義しているが、近年では深層学習に基づいたモデルで研究されることが多い⁶⁾⁷⁾⁸⁾⁹⁾¹⁰⁾。著者らもイングランドの交通データで高速道路を中心とした道路網での短時間交通予測を Graph neural network (GNN)¹¹⁾で試行している¹²⁾。しかし前

報³⁾の検討で比較的良好な結果が得られていること、学習に必要な時間が短い (本研究の範囲では最も説明変数の多いケースでも所要時間は数秒程度だった) ことから本研究でも LightGBM¹³⁾を用い、曜日、時刻や現在、過去の平均速度、通過台数を説明変数として予測を行う AI モデルを作成した。

本研究は前報³⁾の検討に a) 多地点での性能を比較するために渋滞するかしないかの分類予測ではなく速度の回帰予測を行う、b) 前報で対象とした地点の前後地点の予測も行う、c) 前後地点のデータを加えた場合に対象地点の予測精度が改善するか調べる、d) SHapley Additive exPlanations (SHAP)¹⁴⁾により説明変数の重要度を調べる、という改良を加えた。その結果、前報の対象地点の前後では交通の状況が大きく異なっているものの、説明変数をうまく取ることによりある程度の精度で速度予測ができることがわかった。また SHAP による説明変数の重要度で、予測地点の現況だけでなく過去の速度履歴が予測性能の改善に効果があることが明らかになった。一方、予測地点の過去の台数履歴や前後地点の情報は本研究の対象地点では予測性能改善への寄与が小さかった。

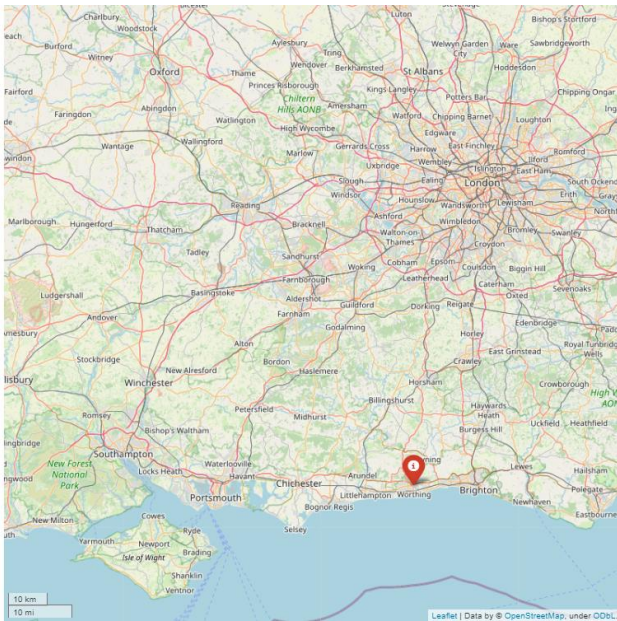


図- 1 本研究で対象とした計測サイト (Id=6743) のイングランド南部での位置。地図は上が北で、右上の道路が環状に取り巻いている都市がロンドン。(©OpenStreetMap contributors)

2. データ

(1) WebTRIS Traffic Flow API

本研究では WebTRIS Traffic Flow API からダウンロードしたデータを使用した。データの詳細については前報を

参照していただきたい。WebTRIS Traffic Flow API のデータには車長別台数、速度別台数の情報もあるが、本研究ではそれらを用いずに時刻と平均速度、平均通過台数の情報だけを使用した。なお元のデータでは速度は mph (mile per hour, 時速マイル) で表されているが、本研究では時速 km (km/h) に換算している。

(2) 本研究で対象とした計測サイト・道路

前報に引き続き、イングランド南部を東西に走行する A27 という一般幹線道路が West Sussex 州の Worthing という街の中心付近を通過する Id=6743 とその前後の測定点を検討対象とした。Wikipedia には「Anundel and Worthing are both areas of known traffic congestion during times of peak usage.」という記述があり¹⁵⁾、イングランド内でも比較的渋滞の発生しやすい地点であると考えられる。その原因としては海と丘に挟まれた地形で、並行して走行する高速道路が無いことなどが考えられる。なお前報では東西両方向の交通流について検討したが、本報では東向きの交通流に的を絞って検討を行った。

図- 2 に対象地点の拡大図を示す。A27 は図の中段付近を東西に走行している道路で、本研究の検討では西側 (図の左側) が上流、東側 (右側) が下流になる。Worthing の街は図の右下で、Worthing 付近で A27 は東西方向に並走する道路 (A2032) と合流している。また A27 の北側は丘陵になっており、近隣に東西方向に並走する道路は存在しない



図- 2 本研究で対象とした計測サイト付近の拡大図。地図は上が北で、中段付近を左右に走行する道路が A27。本研究では西側が上流で東側が下流。青または赤のピンが立っている場所が測定点で 6743 などの番号が Id。右下の赤線が 1 km に対応する (©OpenStreetMap contributors)

本研究では新たに上流側 2 地点 (Id=7728, および 7725) と下流側 1 地点 (Id=7723) も検討対象に加えた。いずれの地点も計測データが揃っている地点としては Id=6743 に最も近い。上流側を 2 地点加えたのに対し下流側が 1 地点なのは下流の Id=7723 が Id=6743 から約 2km と比較的近いのに対し、上流側の測定点が離れていることを補うためと、上流側の方が影響が大きいであろうという予想による。また地図上では A27 に合流したり、並走している番号付きの一般道がいくつか存在するが、WebTRIS Traffic Flow API のデータ中にはこれらの道路に対応する計測地点は存在しない。

(3) 各測定点の QV 図

本研究で対象とした 4 地点での平均速度と通過台数の関係 (QV 図) を図- 3 に示す。各図の右上の数字が地

点 Id を、横軸が各地点での通過台数 (15 分当たりの台数) を、縦軸が平均速度 (km/h) を表している。ただし後述する訓練・評価期間に対応する 2016 年 1 月 1 日から 2019 年 12 月 31 日までのデータを図示している。

本研究で中心的に取り扱う Id=6743 の地点では 15 分当たり通過台数が 350 台付近で QV 図に折れ曲がりが見られ、このあたりに交通容量の限界があると考えられる。同様に約 2 km 下流の Id=7723 でも 15 分当たり 550 台付近に限界が見られる。一方、上流側の Id=7728 と Id=7725 では明瞭な折れ曲がりが見られず、交通容量の限界に達することはほとんど無い。地点ごとに速度の上限は異なっており、Id=7728, および 6743 では時速 50 km, Id=7725, および 7723 では時速 100 km 付近に速度の上限があると考えられる。

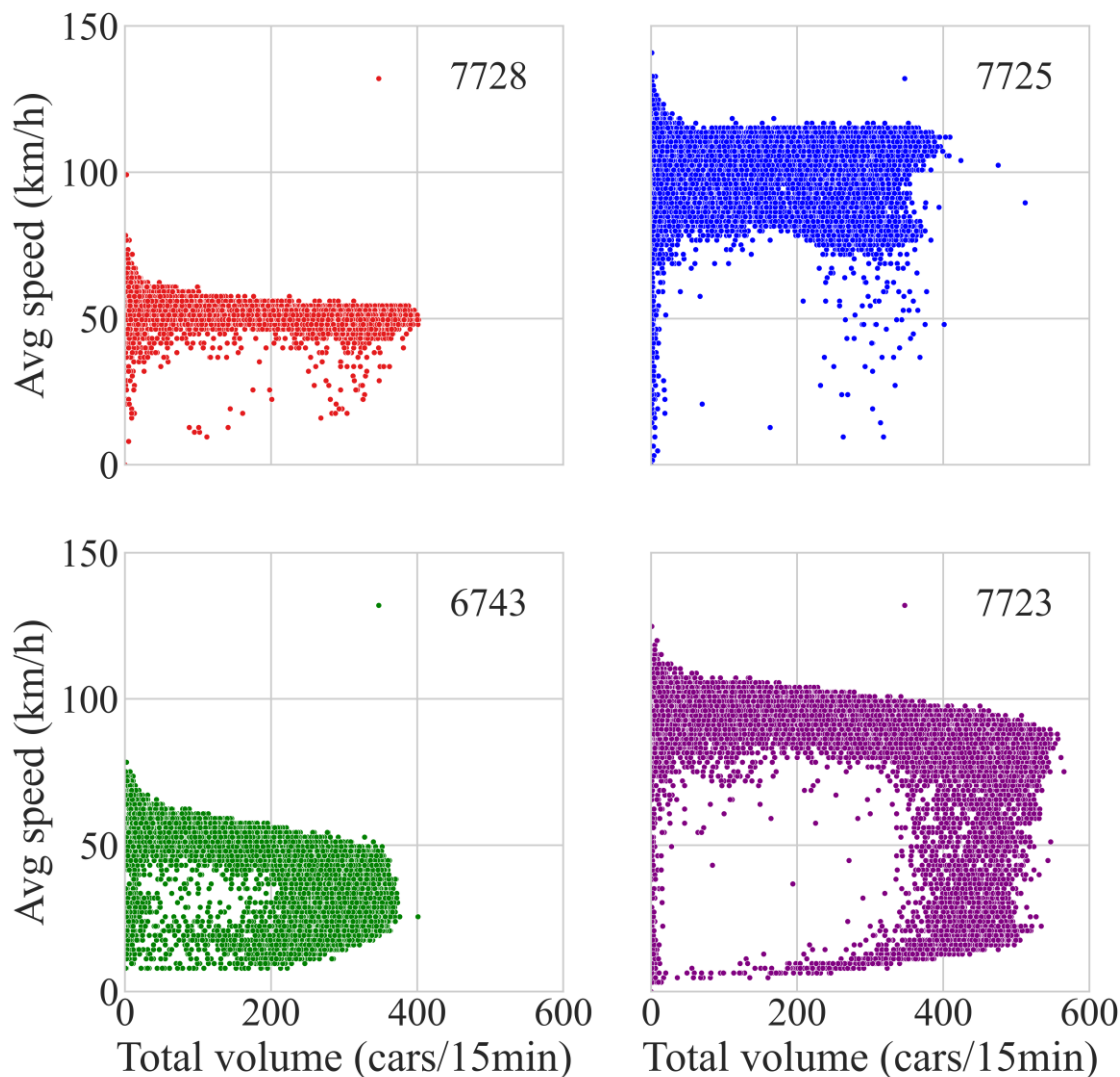


図- 3 各測定点の QV 図の比較。各図の右上の数字は地点 Id を、縦軸は平均速度 (km/h) を、横軸は 15 分当たりの平均通過台数を表す。

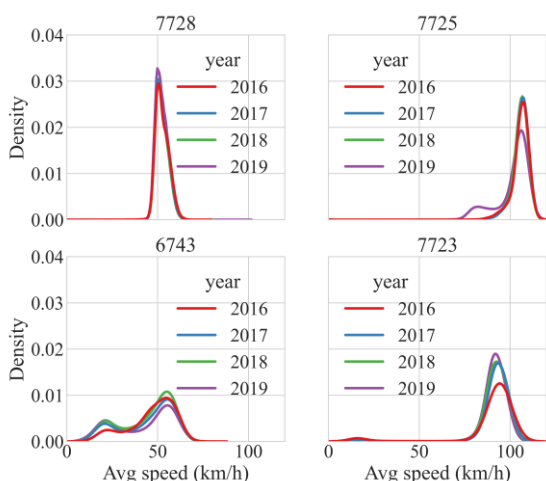


図-4 各測定点の速度分布の時間変化をカーネル密度推定で示す。各図のタイトルは測定点の Id を、凡例は年を示す。また横軸は平均速度（時速 km）で縦軸は相対密度である。

(4) 各測定点の速度分布の時間変化

図-3 の Id=7725 では QV 図の「腕」が他の地点と比較して広い幅を持ち、特に台数の上限付近で2重になっているように見える。この原因を調べるために各地点の速度分布を年ごとにカーネル密度推定で示したのが図-4である。測定地点の Id は各図にタイトルとして表し、1年ごとに集計したカーネル密度推定を色の異なる線で示している。また図の横軸が平均速度を、縦軸が相対密度である。なお通常のヒストグラムではなくカーネル密度推定を用いたのは年ごとの分布の違いを簡潔に示すためである。

図より Id=7725 の地点では 2019 年に速度分布の形状が変化して時速 75 km 付近に小さな第 2 のピークができており、これが QV 図に第 2 の「腕」として表れていると考えられる。同様に Id=6743、および 7723 の地点でも年ごとの速度分布に変化が見られる。これらの時間変化は作成する予測器の性能に悪影響を与えそうであるが、本研究ではそのまま使用することにした。

(5) 訓練・評価データ

説明変数としては交通状況の測定された日時の曜日と時刻、現時点の平均速度、通過台数、60 分前まで遡った過去の平均速度、および通過台数を用い、後述のように組み合わせを変えて複数のモデルを作成した。前報では時刻として「時」だけを用いていたが、本研究では「時」に加えて「分」まで入力変数とした。これによりわずかに性能改善が見られ、予測時間による性能変化が小さくなった。ただしデータの時間分解能が 15 分なので、「分」の取り得る値は 4 通りである。

またダウンロードしたデータの時間分解能が 15 分間隔になるのは 2015 年 4 月以降であり、2020 年以降は

Covid-19 により交通パターンに大きな変化が見られたので 2016 年 1 月から 2018 年 12 月までのデータを訓練データ、2019 年 1 月から 12 月までのデータを評価データとした。全てのモデルにとって公平な比較になるよう、一か所でもデータの欠落している日時のデータは全ての測定地点で取り除いた。

3. モデル

(1) Null (naive) モデル

本研究では現状がそのまま継続すると仮定する null モデルを性能評価の基準とした。前報では「naive モデル」としていたが AI 分野での一般的な呼び方に従った。AI モデルの性能が null モデルよりも良いのは当然という考え方もあるが、短時間交通予測は場所による性能の違いが大きい傾向があるので、AI モデルの満たすべき性能の下限値という位置づけで検討した。

(2) AI 予測器

本研究では AI 予測器として、前報の分類予測で高い性能を示した LightGBM¹³⁾ を回帰予測器として用いた。デフォルトの設定で良好な予測性能を示したため、乱数のシードを固定する以外のオプション変更は行わなかった。また回帰木系の予測器なので曜日や時刻は One-hot エンコーディングせずに数値のまま、速度や台数のデータも正規化せずにそのまま入力した。

(3) AI モデル

表-1 本研究で使用した null モデルと AI モデル。

モデル名	説明変数
Null モデル	(現状が続くと予測)
モデル A	曜日・時刻 (時, および分)
モデル B	モデル A の説明変数, および現時点の平均速度, 通過台数
モデル C	モデル B の説明変数, および 15~60 分前の平均速度, 通過台数
モデル D (Id=6743 のみ)	曜日, 時刻 (時, および分) と Id=7728, 7725, 7723, および 6743 の現時点~60 分前の平均速度, 通過台数

本研究で使用した null モデルと AI モデルを表-1 にまとめると。Null モデルについては前述の通りで、AI モデルの性能基準として用意した。モデル A は曜日と時刻 (時, および分) を説明変数とするモデルで、周期的に速度変化するような状況で高い性能を発揮する。モデル B は曜日と時刻に加え現時点の平均速度と通過台数を説明変数とする。説明変数に現況が加わっている

ので周期的な速度変化に加え、不規則な速度変化にも対応できることが期待される。モデル C ではモデル B の説明変数に加え、60 分前から 15 分前までの平均速度と通過台数を加えた。これは一時間程度の時間をかけて成長するような速度変化のパターンを学習することを想定している。

Id=6743 だけに適用したモデル D は曜日と時刻に加えて本研究で対象とした全 4 地点の 60 分前から現在までの平均速度と通過台数を説明変数としている。これは前後の地点の交通状況が影響するような速度変化を学習することを想定している。

(4) 性能評価

本研究では回帰予測の性能評価指標として二乗平均平方根誤差 (Root Mean Squared Error, 以下 RMSE) を用いた。Max Error のような最悪値を比較する指標と比べると小さな値になりがちで、特に渋滞時には体感より小さい値になりがちだが、誤差を速度と同じ時速 km で評価できるという利点がある。

4. 結果

(1) Null モデル

Null モデルを 4 地点に適用し、予測時間と RMSE の関係を調べたのが図-5 である。地点ごとに色を変えていて、横軸が予測時間 (分) で縦軸が RMSE (km/h) で、下に行くほど性能が良い。いずれの地点でも将来の速度は現在の速度に依存しており、予測時間が短いほど現在の速度との相関が大きいため性能が良くなる (RMSE が小さい)。図-4 の速度分布で分布の広がりが小さく、速度変化が小さいと考えられる Id=7728 で RMSE 値が最も小さく、予測時間による変化も小さかった。

QV、速度分布が Id=7728 と同様の傾向を示している Id=7725 がそれに続き、QV 図で明瞭な折れ曲がりが見られた Id=6743、および 7723 では予測時間が長くなるほど予測性能が悪化する傾向が見られた。

(2) モデル A

曜日と時刻を説明変数とするモデル A を 4 地点に適用し、図-5 と同様に予測時間と RMSE の関係を調べたのが図-6 である。このモデルでは現在の状態は一切使わずに予測を行っているため、RMSE の予測時間への依存性はほとんど見られない。説明変数の時刻を「時」の単位にした場合には予測時間によるわずかな性能変化が見られるが、「分」まで加えて学習した予測器では性能がほぼ一定になる。

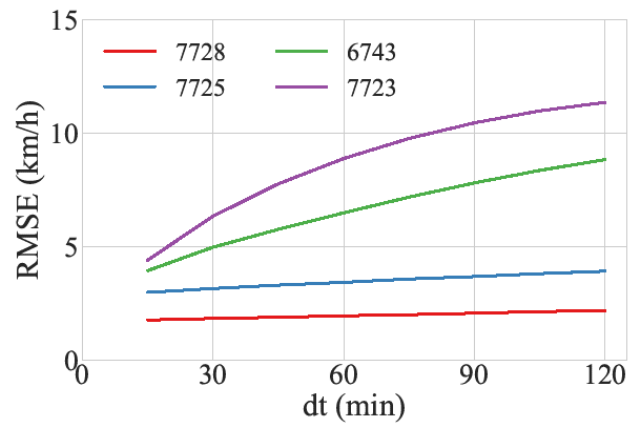


図-5 Null モデルの RMSE の比較。横軸は予測時間 (分) で縦軸が RMSE (km/h)。

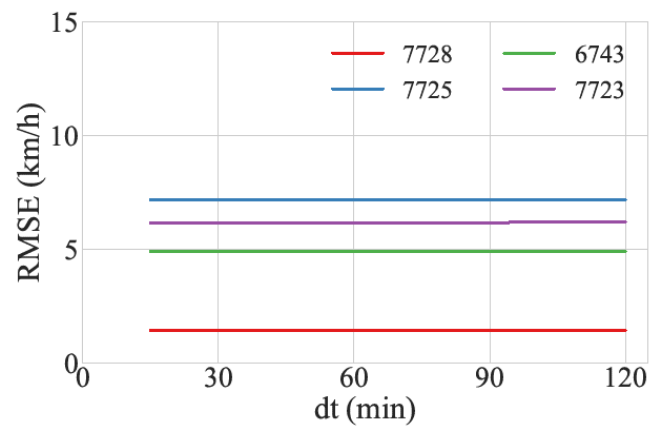


図-6 モデル A の RMSE の比較。横軸は予測時間 (分) で縦軸が RMSE (km/h)。

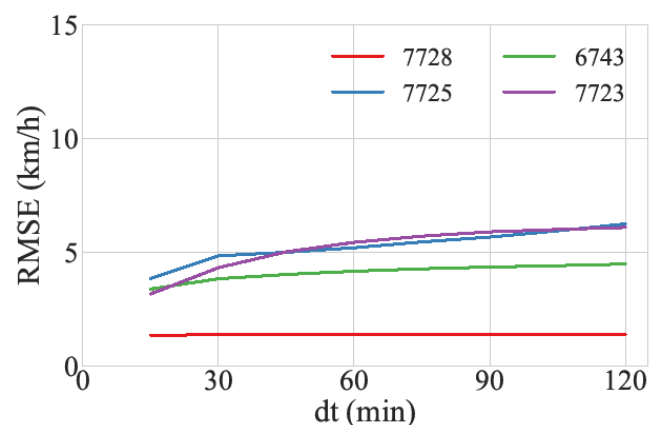


図-7 モデル B の RMSE の比較。横軸は予測時間 (分) で縦軸が RMSE (km/h)。

図-3 に見られるように平均速度が比較的遅く、速度変化も小さい Id=7728 では平均速度に近い値を予測することにより RMSE も小さくなる。

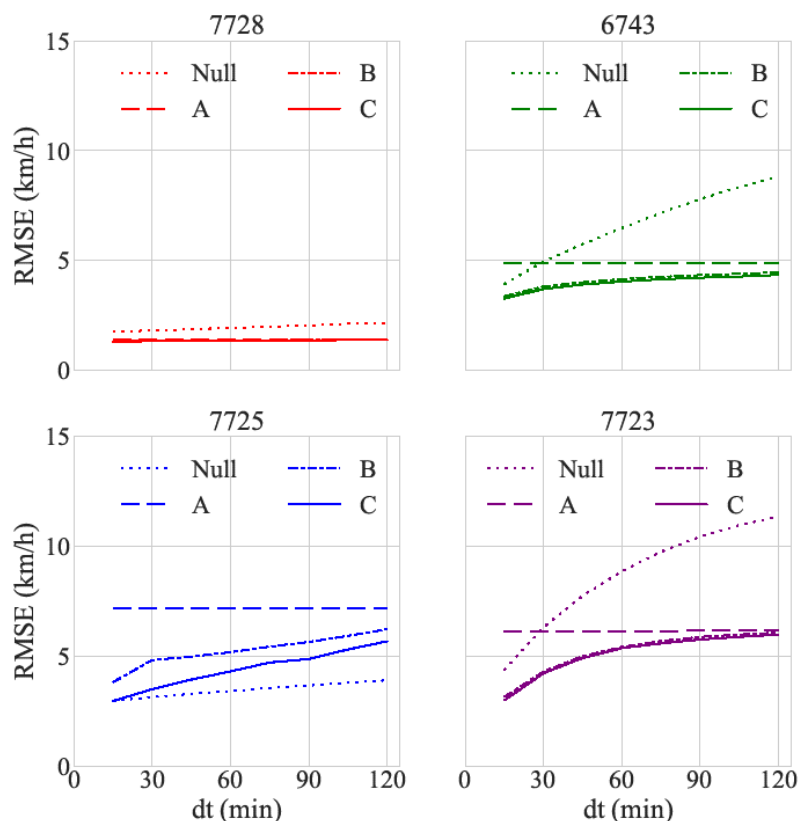


図- 8 各測定地点での null モデル，およびモデル A~C の性能の比較. 各図のタイトルは測定点の Id を，横軸は予測時間 (分)，縦軸は RMSE (km/h) を表す. また破線は null モデルの，破線はモデル A の，一点鎖線はモデル B の，そして実線はモデル C の結果を示す.

(3) モデル B

曜日と時刻に加えて現在の平均速度と通過台数を説明変数とするモデル B の RMSE を図- 5 と同様に示したのが図- 7 (前ページ) である。このモデルでは現況を説明変数として利用しているため、どの地点でも予測時間が短いほど性能が良くなっている。

(4) モデル C

モデル C では曜日と時刻，現在の平均速度と通過台数に加えて 60 分~15 分前の平均台数と通過台数を説明変数とした。本研究の各測定地点での null モデル，およびモデル A~C の性能の比較を図- 8 に示す。各測定点ごとにグラフを作成し、それぞれ Id をタイトルに示している。また横軸は予測時間 (分) を，縦軸は RMSE (km/h) で下に行くほど性能が高い。点線が null モデルの，破線がモデル A の，一点鎖線がモデル B の，そして実線がモデル C の結果を示している。Id=7725 以外の地点ではモデル B とモデル C の値は近く、この図のスケールではほとんど重なって見える。

モデル A の性能が高い Id=7728 では説明変数を加えることによる性能向上は小さいが、他の地点ではモデル B、

および C の性能はモデル A から改善している。また Id=7725 を除く 3 地点では null モデルの性能をモデル B、C の性能が上回っており、AI による予測が機能していると言える。Id=7725 についても 15 分後では null モデルとモデル C の性能は同程度になっており、AI の予測は完全な外れとは言えない。

図- 8 のスケールではモデル B と C の違いはほとんど見られなかったが、Id=6743 について詳細にプロットした図- 9 (次ページ、詳細は後述) で比較するとモデル C の性能はモデル B よりわずかに良くなっている。これは過去の交通状況にも予測性能改善に寄与する要素が存在するということである。どの要素に予測性能改善の効果があるのかは後の節で分析する。

Null モデルの性能が最も良かった Id=7725 の予測例を図- 10 (次ページ) に示す。図は評価期間 (2019 年) で Id=7725 の地点で大きな速度変化が見られた期間の平均速度の変化 (灰色線) と null モデルによる 60 分後の予測 (平均速度を 60 分間後方に移動したもの、赤線)、モデル C による 15 分予測、および 60 分予測 (それぞれ青線、緑線) を比較している。図の横軸は日時 (日+時) を，縦軸は平均速度 (km/h) を示している。

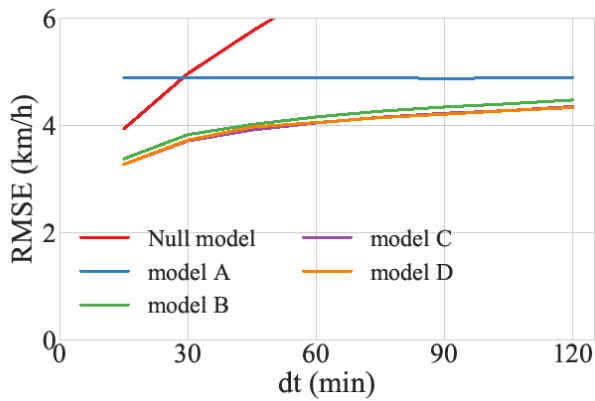


図- 9 Id=6743の地点での各モデルのRMSEの比較。横軸は予測時間(分)で縦軸はRMSE(km/h)。縦軸方向に拡大したため、nullモデルは途中でまてになっている。

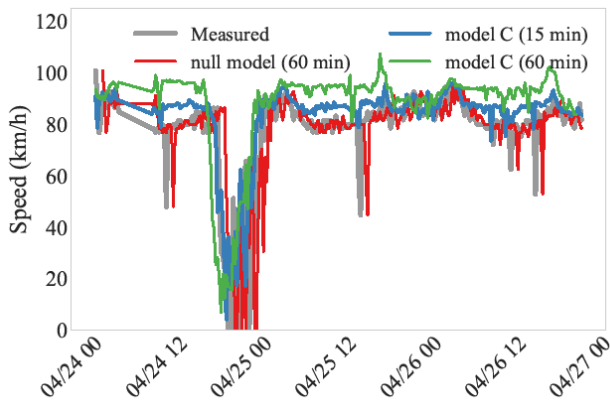


図- 10 Id=7725の地点での平均速度の時系列変化と予測例。横軸は日時(2019年の日付と時刻)、縦軸は平均速度(km/h)。灰色線が実際の値で、赤線がnullモデルによる60分の、青線と緑線がモデルCによるそれぞれ15分、60分の予測。

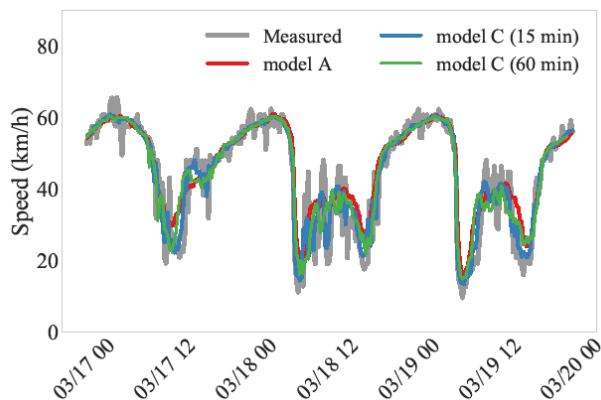


図- 11 Id=6743の地点での平均速度の時系列変化と予測例。横軸は日時(2019年の日付と時刻)、縦軸は平均速度(km/h)。灰色線が実際の値で、赤線がモデルAによる15分前の、青線と緑線がモデルCによるそれぞれ15分前、60分前の予測。

図より null モデルは遅れがあるものの速度変化の傾向

に追従できているのに対し、モデルCは現在の速度が継続されると予測する傾向があることがわかる。この地点で通常は速度が安定しているため、AIも小さな変化があったとしても元に戻ると予測する傾向があるためであると考えられる。また図- 4の訓練データと評価データの速度分布の違いも原因となっている可能性がある。

同様にId=6743の地点について実際の速度とAIによる予測を比較したのが図- 11である。横軸は2019年の日時を、縦軸は平均速度(km/h)を表している。灰色線は実際に測定された速度を、赤線はモデルAによる15分予測を、青線と緑線はそれぞれモデルCによる15分予測と60分予測である。モデルAは予測時間による性能の違いが小さいため、120分前の予測も同様になる。またこの年の3月17日は日曜日で、昼間の時間帯の平均速度の落ち込みは小さいが、翌日以降は昼間、特に朝と夕方の通勤時間帯に大きな速度の落ち込みが見られる。

この地点の速度変化は周期性が高く、曜日、日時だけを説明変数としたモデルAでも速度変化の傾向をよく予測できている。一方、日曜日の日中の速度変化や、平日昼間の速度変化は過去60分間の交通状況を説明変数にしたモデルC、特に15分前の予測が最も良く傾向をとらえている。モデルCは120分前の予測でも決定係数が0.8以上になっており、この地点に関しては高い予測性能を持っていると言える。

(5) モデルD

モデルDではId=6743の地点の予測精度を改善するために曜日、日時と図- 2に示した4地点の60分前～現在の平均台数と通過台数を説明変数として訓練を行った。各モデルの予測性能を比較した結果を図- 9に示す。図の横軸は予測時間(分)を、縦軸はRMSE(km/h)を示している。モデルBとCの違いが分かりやすくなるように縦軸を拡大表示したため、nullモデルのプロットが途中で切れている。全体像は図- 8を参照されたい。

モデルBと比較するとモデルC、およびDはRMSE値がわずかに改善されているが、モデルCとDの差は非常に小さい。従ってこの地点については前後地点のデータを入力することによる予測性能の改善は見られないと言える。

(6) 説明変数の重要度の検討

次にどの説明変数が予測にとって重要かの検討を行った。重要度を求めるための手段はいくつかあるが、本研究では今後、個別の予測への説明変数の寄与も検討することを視野に入れてSHAPを使用した。また図- 9でモデルCとモデルDの性能の差がほとんど見られなかったことからId=6743の地点の予測を行うモデルCを対象を絞って検討を行った。

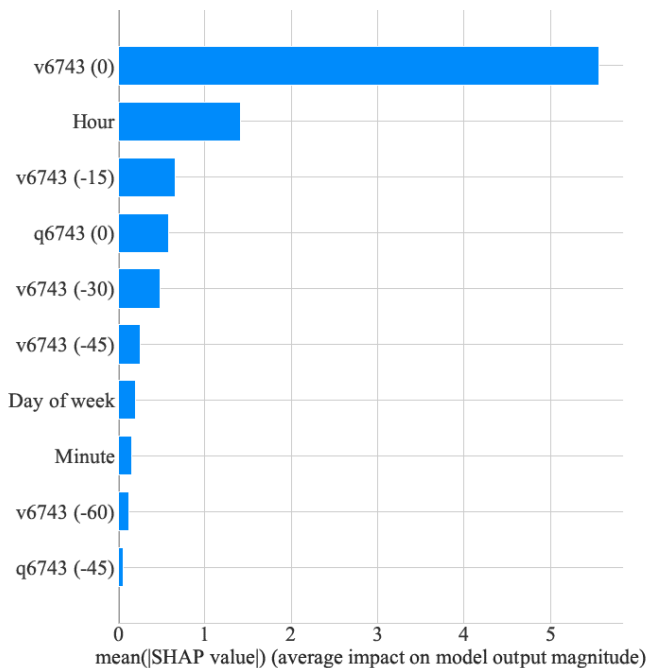


図- 12 Id=6743 の地点の 15 分後の速度を予測するモデル C の予測器の SHAP 値。縦軸は説明変数名で横軸は SHAP 値の絶対値の平均。上位 10 変数の結果を示す。

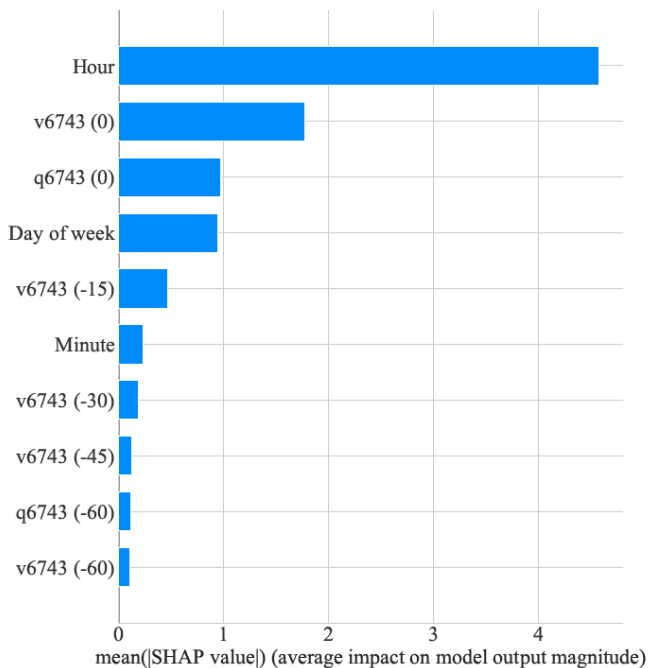


図- 13 Id=6743 の地点の 120 分後の速度を予測するモデル C の予測器の SHAP 値。縦軸は説明変数名で横軸は SHAP 値の絶対値の平均。上位 10 変数の結果を示す。

Id=6743 の 15 分後の速度予測を行う予測器に SHAP を適用した結果を図- 12 に示す。図の縦軸は変数名で、「Day of week」は曜日、「Hour」、および「Minute」はそれぞれ時、分を、v6743 (0) は現在の平均速度、v6743 (-15) は 15 分前の平均速度、q6743 (0) は現在の通過台数、などを表している。ここでは上位 10 変数の

結果を取り出して示している。横軸（棒グラフの棒の長さ）は SHAP 値の絶対値の平均を表していて、この値が大きいくほど説明変数の重要度が大きい。SHAP 値は個別の予測に対して求められる上正負の値を取るため、それぞれの説明変数について SHAP 値の分布を調べることができる。SHAP 値の分布まで使うことで説明変数の効果を詳細に検討できるが、ここでは単純化のため絶対値の平均を用いている。本研究の範囲ではこのような使い方も説明変数の重要度を十分良くとらえられている。

図より 15 分後の予測では v6743 (0)、つまり現在の平均速度が最も重要であることがわかる。「時」、15 分前の平均速度、現在の通過台数、30 分前、45 分前の平均速度、曜日、「分」と続く。交通状況を表す変数の中では通過台数よりも平均速度の方が速度予測には重要であることがわかる。

図- 12 の上位の説明変数がどの程度予想に寄与しているか調べるために、上位 8 個（「Minute」まで）の説明変数だけを使用する予測器を作成、訓練し、評価データに適用したところ、RMSE 値は 3.27 km/h だった。この値はモデル C とほぼ同じだった。一方、上位 7 個（「Day of week」まで）の説明変数を使用した予測器の場合は 3.33 km/h となり、わずかに予測性能が悪化した。従ってこの地点で速度予測を行うために重要な説明変数は 8 個（曜日、日時、45 分前から現在までの平均速度、現在の通過台数）であると言える。

同様に 120 分後の予測を行う予測器に SHAP を適用した結果を図- 13 に示す。図の見方は図- 12 と同様である。15 分後の予測を行う予測器と比較すると、現在の速度 (v6743(0)) の重要度が小さくなり、「Hour」が最も重要になっている。モデル C で 120 分後の予測を行う予測器の RMSE 値が 4.35 km/h だったのに対し、図- 13 の上位 8 個 (v6743(45)まで) の説明変数を使用した予測器の RMSE 値は 4.37 km/h、上位 10 個 (v6743(-60)まで) では 4.36 km/h であった。120 分後の予測を行う場合でも過去の速度履歴は重要で、過去の通過台数もわずかに有用な場合があると言える。

5. 結論

本研究では交通のデジタルツインの適用例として、イングランドの交通データに対して短時間速度予測を行った。対象とした地点は前報と同じく Worthing 近郊の一般道だが、より渋滞頻度の高い東方向に限定し、新たに前後地点の検討と、前後地点のデータを加えた場合の検討も行った。また交通状況の異なる複数地点を取り扱うことから、渋滞するかしないかの分類予測ではなく、直接速度を予測する回帰予測を行い、主に RMSE による性能

評価を行った。

予測モデルには前報で性能が良好だった LightGBM を用い、説明変数には曜日、時刻（時分）と過去から現在までの平均速度、通過台数の履歴などから組み合わせを変えていくつかのモデルを作成した。また性能の比較対象として現状が続くと仮定する null モデルも用意した。

訓練データと評価データで平均速度の分布に違いが見られた 1 地点を除く 3 地点で曜日、時刻、現況を説明変数としたモデルの性能が null モデルを上回り、AI による予測効果が確認できた。予測誤差を RMSE で表すと 5 km/h 前後で比較的良好であると言える。ただし RMSE は全期間を二乗平均して計算する指標であり、局所的な予測性能の悪化は評価できない。今後はどのような状況で予測誤差が大きくなるのかなど、個別ケースでの評価が必要であると考ええる。

予測性能は説明変数を曜日と日時、現況としたモデルでおおむね頭打ちになり、それに過去の履歴を加えるとわずかに性能が改善した。SHAP で説明変数の重要度を検討したところ、過去の実績履歴に予測の改善効果があることがわかった。SHAP は個々の予測に対して説明変数の寄与度を求めることができるので、今後、どのような交通状況の時に過去の履歴が予測を改善するのか検討を進める。

前報で対象とした地点については前後の地点のデータも説明変数に加えて速度予測を行ったが、予測性能の改善は見られなかった。これについては時間分解能が 15 分と粗いことと、測定点の間隔が 2~10 km 程度と時間分解能に対して大きいこと、一般道なので測定点同士の間に分岐があることなどが原因として考えられる。一方、著者らのグループが測定点の密度が高い高速道路網で Graph Neural Network (GNN) を用いた予測を行った場合には LightGBM に対して予測性能が改善するケースが見られた¹²⁾ので、全ての場合で周辺地点のデータに予測性能の改善効果が無いわけではない。今後どのような場合に周辺地点のデータが予測性能を改善するのか検討を進めたい。

同じ検討では GNN による予測が全てのケースで LightGBM に対して優れた性能を示しているわけではない。対象地点によって予測モデルの得手不得手があるものと考えられる。また予測に必要な計算コストは GNN では LightGBM より高いなど、予測モデルによって大きな違いがある。今後は ARIMA モデルなど AI 以外のモデルも含めた複数手法を多くの地点に適用して、どのような地点でどのような予測モデルを使用するのが適当か、

という検討を行う必要があると考える。

参考文献

- 1) highways england : <https://webtris.highwaysengland.co.uk/>.
- 2) WebTRIS Traffic Flow API : <https://webtris.highwaysengland.co.uk/api/swagger/ui/index>.
- 3) 宮崎利行, 大沢明瑠, 菊池恵和, 菅原宏明 : イングランドの交通オープンデータと AutoML による短時間渋滞予測, AI・データサイエンス論文集, 3 巻, J2 号, pp. 268-276, 2022.
- 4) 国土交通省 道路局:道路システムの DX ~xROAD の推進~ : <https://www5.cao.go.jp/keizai-shimon/kaigi/special/reform/wg6/20220419/pdf/shiryou1-2.pdf> (2022 年 12 月参照).
- 5) E. I. Vlahogianni, M. G. Karlaftis, J. C. Golias : Short-term traffic forecasting: Where we are and where we're going, Transportation Research Part C: Emerging Technologies, Volume 43, Part 1, pp.3-19, 2014.
- 6) N. G. Polson, V. O. Sokolov : Deep learning for short-term traffic flow prediction, Transportation Research Part C: Emerging Technologies, 79, pp.1-17, 2017.
- 7) Z. Zheng, W. Chen, X. Wu, P. C. Y. Chen, J. Liu : LSTM network: a deep learning approach for short-term traffic forecast, IET Intelligent Transport Systems, Vol.11 Iss.2, pp.68-75, 2017.
- 8) B. Yu, H. Yin, Z. Zhu : Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting, IJCAI, 2018.
- 9) 坪田隆宏, 砂侑奈, 吉井稔雄 : 畳み込みニューラルネットワークによる都市間高速道路の渋滞発生予測モデル, 第 66 回土木計画学研究発表会・講演集, 2022.
- 10) 田子祐亮, 中西雅一, 菅芳樹, 西田匡志, 山下和彦, 日下部貴彦 : 都市間高速道路の交通状態を考慮した LSTM による旅行時間予測手法, 第 66 回土木計画学研究発表会・講演集, 2022.
- 11) J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, M. Sun : Graph neural networks: A review of methods and applications, AI Open, Vol.1, pp. 57-81, 2020.
- 12) 緒方陸, 宮崎利行, 菊池恵和, 村野祐太郎, 菅原宏明 : 短時間走行速度予測における機械学習手法と Graph Neural Network の比較, AI・データサイエンス論文集, (投稿中) .
- 13) LightGBM : <https://lightgbm.readthedocs.io/en/latest/>.
- 14) S. M. Lundberg, and Su-In Lee : A unified approach to interpreting model predictions, Advances in Neural Information Processing Systems, 2017.
- 15) https://en.wikipedia.org/wiki/A27_road (2022 年 12 月閲覧) .
- 16) Scikit-learn : <https://scikit-learn.org/stable/index.html>.

(Received December 31, 2022)

(Accepted January 31, 2023)

AI SHORT-TERM SPEED PREDICTION ON AN ARTERIAL ROAD OF ENGLAND

Toshiyuki MIYAZAKI, Riku OGATA, Yutaro MURANO, Yoshikazu KIKUCHI, and Hiroaki SUGAWARA

Short-term speed prediction was investigated as a way to utilize the digital twin of traffic conditions. Traffic data for England is publicly available in real time, and its historical data can be downloaded as open data. In this study, we selected a relatively congested point of an arterial road from the downloaded data, compared the traffic condition data with neighboring points of the target, and predicted the speed using machine learning (AI). In addition, we compared the importance of features using SHapley Additive ex-Planations (SHAP), and found that not only the current situation, but also the past speed history had an effect on improving the prediction performance. In the authors' previous study, only the data of the forecast point were used as input variables for the short-term forecasting, but in this study, the traffic conditions of the neighboring points were input to see if there was any improvement in forecasting performance. As a result, no improvement in forecasting performance was observed even when the traffic conditions at the neighboring points were added as input variables.