将来予測精度を考慮した舗装劣化予測モデル の精度評価手法の提案と検証

奥田 知之¹・鈴木 康豊²・神武 直彦³

¹正会員 慶應義塾大学院 システムデザイン・マネジメント研究科
(〒223-8526 横浜市港北区日吉4-1-1協生館)
E-mail:t.okuda@keio.jp

²株式会社パスコ 中央事業部 技術センター インフラマネジメント部
(〒153-0043 東京都目黒区東山1-1-2 東山ビル5階)

³慶應義塾大学院 システムデザイン・マネジメント研究科
(〒223-8526 横浜市港北区日吉4-1-1協生館)

効率的な舗装維持管理を実現する為に、舗装劣化予測により将来の状態変化を予測し、点検、補修、更新等の維持管理コストを最適化する必要がある。様々な舗装劣化予測モデルが提案されているが、それぞれのモデルがどのような状況下で最適なのかを判断する事は容易ではない。

そこで、本研究は時刻情報を持つ点検履歴データに対する、より実運用に近い予測モデルの予測精度を求める事が可能なタイムスライスクロスバリデーションを提案する。提案手法は、時刻方向に点検データを分割し、学習データからモデルのパラメータ推定を行い、そのパラメータを用いて試験データの予測を行った結果から、現実的な予測精度評価を行う。提案手法をRNNモデルによる舗装劣化予測に適用し、従来手法より現実的な将来の予測精度を評価可能であることを検証した。

Key Words : *pavement management, pavement condition survey, model evaluation, cross-validation, recurrent neural network*

1. はじめに

近年、社会資本の維持管理を効率的に行っていく事が求められている。舗装の維持管理は、従来よりPMS(Pavement Management System)による効率的な維持管理計画策定支援が研究、開発されている。PMSの主要な機能として舗装劣化予測があり、舗装管理者は劣化予測結果を元に効率的な維持管理計画を策定する。

舗装維持管理計画は長期的に管理対象路線全体として望ましい水準を維持するための長期的な維持管理(ネットワークレベル)の計画と、舗装の単位区間ごとの補修・更新・点検等の短期的な維持管理(プロジェクトレベル)に分類される。プロジェクトレベルの精度の高い予測が可能になれば、予測が容易な区間は点検頻度を下げる事や、破壊が深刻化する前に補修を行う事で長寿命化する予防保全を実現し、他方でまだ補修が必要ない箇所の補修を先送りする事、管路工事等と同時に舗装工事を実施する事¹⁾等により、維持管理コスト削減が可能になる。しかし現状は維持管理計画で管理基準を設け、定期点検及び日常点検によりその基準を下回った事が測定

された後、補修工事の計画を立てる対処療法的な維持管理が行われている²⁾。その理由としては、舗装の単位区間に対する真の値を予測する事は困難である事が上げられる。まず、舗装の状態を表す路面性状値等の点検結果自体に、点検担当者や装置の違い、天候の違い等による系統誤差が含まれており、真の値の推定が不可能である。それに加え、舗装の素材の変化や、新規・廃止路線の影響で大幅に交通量が変化する等の未来の不確定な事象は考慮する事が出来ない。しかし、予測精度を精確に求める事ができれば、その精度範囲内のばらつきを許容する事で、単位区間に対する予測を用いた維持管理の効率化が可能になる。

また、従来の舗装劣化予測では、定量的な予測精度評価が行われていないか、それぞれ異なる方法により評価されている。そのような場合、様々な劣化予測モデルのどの方法が最適か、判断が難しいという課題がある。

本研究はこれらの点について着目し、舗装の単位区間に対する現実的な予測精度を評価する手法を提案する。

2. 舗装劣化予測モデルの予測精度評価方法

(1) 舗装劣化予測モデルとその評価方法

従来、マルコフチェーンモデル^{3,4)}や重回帰モデル⁵⁾による舗装劣化予測モデルは、パフォーマンスモデルとして構築されている。パフォーマンスモデルは、全てのデータを学習データとしてモデルのパラメータ推定(学習)に用いて、大型車交通量か、補修又は新設後の経過時間に対して、評価単位区間を任意にグループ化したグループ全体の平均的な状態変化(パフォーマンスカーブ)を示すモデルである⁹⁾。そのため舗装の単位区間の予測には適さないが、グループ間の違いや、グループの平均的な変化の予測、中長期的な予測、説明変数の寄与の程度を把握する事に用いられる。

それに対し多くの説明変数とパラメータを用いた、複雑で自由度の高いニューラルネットワーク(NN: Neural Network)のようなモデルでは、単位区間の予測を行う事が可能である。しかし、学習に使用したデータの再現は容易である一方、学習データに過適合する事で、未知のデータの予測精度が低下する過学習という問題がある。過学習を検知・抑制する為に、学習データと試験データを分けるクロスバリデーション(Cross-Validation)による予測精度評価が行われる。NNモデルの最も基本的な構造であるマルチレイヤパーセプトロン(MLP: Multi-Layer Perceptron)により、わだち掘れのパフォーマンスモデルを構築した事例^{6,7)}及び、舗装の点検データを時系列データとして扱い、時系列をモデル化できる再帰型ニューラルネットワーク(RNN: Recurrent Neural Network)モデルにより、アメリカで用いられるPSI(Present Serviceability Index)の予測精度を、ランダムにデータを分割するクロスバリデーションにより求めた研究⁸⁾がある。また、筆者らも既にRNNモデルによるわだち掘れ量の予測モデルを提案しており、重回帰モデルよりも予測精度が良い事を示している⁹⁾。

前者のパフォーマンスモデルは、全てのデータを学習に用いて、学習データの平均的なパフォーマンスを表現できる事に重点が置かれているため、学習に用いていない未知のデータに対する予測精度の評価は行われない。また、RNN等の時系列予測モデルが主に適用されている文字データや音声データが持つ時系列情報は、開始から終了までの相対的な時系列である。従って複数のデータ間が完全に独立しているため、学習データと試験データをランダムに分割する通常のクロスバリデーションで予測精度の検証が可能である。それに対し、本研究で対象とするような点検データの時系列情報は、全てのデータが現実の時間という絶対的な時系列情報を持つ。そのため、時系列点検データの予測モデルを実運用で用いる際には、ある時刻までのデータを学習し、そこまでの情

報だけを元に、それ以降の時刻を予測する。つまり、実運用上は予測を行う時刻の情報は一切学習していない状態で予測を行う。しかし、絶対的な時刻情報を持つ点検データでは、文字データや音声データのように通常のクロスバリデーションによりランダムに学習データと試験データに分割すると、精度評価を行いたい、予測する時刻の情報も学習する事になってしまう。

また、点検データの予測精度評価が劣化予測モデルにより異なると、どのような場合にどの劣化予測モデルが最適であるのか、相互比較が不可能であるという課題がある。それらの事から、舗装の単位区間に対する点検データ履歴を時系列データと捉え、モデル横断的に評価可能で、かつ現実的な将来の予測精度を評価する方法が必要である。

以上の問題認識から、本研究は学習データと試験データを時系列方向に分離する手法により、点検データに対する現実的な予測精度を評価するタイムスライスクロスバリデーション手法を提案する。

3. タイムスライスクロスバリデーション手法の提案

本研究は時系列データである舗装の単位区間に対する劣化予測モデルに対する、クロスバリデーション手法を提案する。まず検証に用いる、NNモデルを時系列データに適用できるように拡張したRNNモデルの概要を説明し、次に提案するタイムスライスクロスバリデーションの説明を行う。

(1) RNNモデル

a) MLPモデル

NNモデルの最も基本的な構造は、図-1に示す生物の神経間の接続を簡易にモデル化したニューロンを多層に接続したMLPモデルである。 n 層目 j 番目のニューロンの状態を表す活性 $a_j^{(n)}$ は次式により求める。

$$a_j^{(n)} = \sum_i \omega_{ji}^{(n)} z_i^{(n-1)} \quad (1)$$

ここで、 $\omega_{ji}^{(n)}$ は入力層から数えて $(n-1)$ 層($n=1, \dots, N$)の i 番目のニューロンから n 層の j 番目のニューロンへの結合荷重パラメータであり、 $a_j^{(n)}$ について活性化関数 $g(x) = \max(0, x)$ で変換した

$$z_j^{(n)} = g(a_j^{(n)}) \quad (2)$$

は、 $a_j^{(n)}$ のニューロンの出力であり、次層のニューロンの入力の一つとなる。

(1)式と(2)式を $z_i^{(n-1)} = x_i$ である入力層から、 $a_j^{(n)} = y_i$

図-1 NN の概念図

図-2 RNN の概念図

である出力層まで繰り返す事で、予測対象である出力値 y_1 を求める計算は、順伝播計算と称する。

b) GRUモデル

RNN モデルは、MLP に内部状態 $\mathbf{h}_t^{(n)}$ を追加し、 $\mathbf{h}_t^{(n)}$ を次の時刻へ伝播するための時刻方向の接続を追加したモデルである。RNN の概念図を図-2 に示す。図中の赤線で示された縦方向の接続が時刻 $t-1$ から、 t の自分自身へ内部状態を伝える再帰的な接続を示す。

様々な手法がある RNN モデルの中でも、学習は早く精度は従来の RNN モデルと同等という研究があることから、本研究では GRU(Gated Recurrent Unit)モデルを用いた¹⁰⁾。時刻 t における $n-1$ 層目の出力を列方向に全て並べた行列を $\mathbf{z}_t^{(n-1)}$ とすると、GRU の内部状態 $\mathbf{h}_t^{(n)}$ は次のように求める。

$$\begin{pmatrix} \mathbf{u}_t^{(n)} \\ \mathbf{r}_t^{(n)} \end{pmatrix} = \begin{pmatrix} \sigma(\mathbf{W}_u^{(n)} \mathbf{z}_t^{(n-1)} + \mathbf{U}_u^{(n)} \mathbf{h}_{t-1}^{(n)}) \\ \sigma(\mathbf{W}_r^{(n)} \mathbf{z}_t^{(n-1)} + \mathbf{U}_r^{(n)} \mathbf{h}_{t-1}^{(n)}) \end{pmatrix} \quad (3)$$

$$\tilde{\mathbf{h}}_t^{(n)} = \tanh(\mathbf{W}_h^{(n)} \mathbf{z}_t^{(n-1)} + \mathbf{U}_h^{(n)} (\mathbf{r}_t^{(n)} \circ \mathbf{h}_{t-1}^{(n)})) \quad (4)$$

$$\mathbf{h}_t^{(n)} = (1 - \mathbf{u}_t^{(n)}) \mathbf{h}_{t-1}^{(n)} + \mathbf{u}_t^{(n)} \tilde{\mathbf{h}}_t^{(n)} \quad (5)$$

ここで $\sigma(\cdot)$ はシグモイド関数、 $\circ$ はアマダール積を表す。 $\mathbf{W}_u^{(n)}, \mathbf{W}_r^{(n)}, \mathbf{W}_h^{(n)}$ は同じ時刻内の、 $\mathbf{U}_u^{(n)}, \mathbf{U}_r^{(n)}, \mathbf{U}_h^{(n)}$ は時刻間の結合荷重パラメータ行列である。GRU は式

(5)で求める内部状態 $\mathbf{h}_t^{(n)}$ を次の層と次の時刻の自分自身への入力とする。これ以降、GRUを用いた RNN モデルの事を RNN モデルと称する。

c) GRUモデルのドロップアウト

GRU モデルに Semenita ら¹¹⁾のドロップアウトを適用した。ドロップアウトは NN モデルの過学習を抑えて予測精度を向上させる、シンプルだが効果的な手法であり、ドロップアウト率 p に応じて、学習時のみランダムにニューロンを無効化し、強制的にモデルの自由度を抑制する事で過学習を抑える。予測を行う際にはニューロンを無効化せずに、 $1/p$ 倍にニューロンが増えていることを補償する為、各層から次の層への出力を全て p 倍する。なお、この p のような値はモデルの結合荷重パラメータを推定するためのパラメータという意味で、ハイパパラメータと称す。

(2) タイムスライスクロスバリデーション手法

本研究で提案するタイムスライスクロスバリデーション手法は、舗装点検データの特徴性を考慮して、舗装劣化予測モデルの予測精度を求める際に必要なデータの分割方法を決定し、それを用いてクロスバリデーションを行う手法を提案するものである。従って予測する舗装点検データと予測モデルの種類は問わない事から、用いるデータを $\mathbf{D}$ 、予測モデルを $\mathbf{y}(\mathbf{x})$ と抽象化して説明を行う。

a) パフォーマンスモデル

共通する時刻情報 $t (t=1,2,\dots,T)$ を持つ点検結果データの時系列を列方向に、舗装の単位区間 $i (i=1,2,\dots,I)$ を行方向に持つデータを全て並べた入力データ行列

$$\mathbf{D} = \begin{bmatrix} \mathbf{x}_{1,1} & \mathbf{x}_{1,2} & \dots & \mathbf{x}_{1,T} \\ \mathbf{x}_{2,1} & \mathbf{x}_{2,2} & \dots & \mathbf{x}_{2,T} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{I,1} & \mathbf{x}_{I,2} & \dots & \mathbf{x}_{I,T} \end{bmatrix} \quad (6)$$

があるとする。実際の点検データは、例えば途中から路線の延伸により増加した箇所はそれ以前の情報が無い等、全ての箇所で共通する時刻情報を持つわけではないが、ここでは説明の便宜上このように示す。パフォーマンスモデルは、 $\mathbf{D}$ 全体を用いてモデルの学習を行う。

$\mathbf{D}$ を学習する事で得たモデルの推定パラメータ(NN では結合荷重パラメータ)を並べたベクトルを $\mathbf{w}_D$ とする。用いるモデルを予測関数として $\mathbf{y}(\mathbf{x})$ で表すと、パフォーマンスモデルは $\mathbf{y}(\mathbf{D}, \mathbf{w}_D)$ と書ける。これは、 $\mathbf{w}_D$ の推定に $\mathbf{D}$ を使い、パフォーマンスの予測精度評価も $\mathbf{D}$ で行う事を表し、後述する過学習の抑制と現実的な予測精度の評価が不可能である。

b) 従来手法のクロスバリデーションによるハイパパラメータの探索と過学習の検証

図-3 RNNモデルのハイパパラメータ決定フロー

それに対し、NN モデルで行われる一般的なクロスバリデーション(従来手法)では、次のように行方向(単位区間方向)にデータを分割する。 $r\%$ を試験データ $\mathbf{D}_T$ 、残りを学習データ $\mathbf{D}_L$ とすると、

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_T \\ \mathbf{D}_L \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{1,1} & \mathbf{x}_{1,2} & \cdots & \mathbf{x}_{1,T} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{x}_{l,1} & \mathbf{x}_{l,2} & \cdots & \mathbf{x}_{l,T} \\ \hline \mathbf{x}_{l+1,1} & \mathbf{x}_{l+1,2} & \cdots & \mathbf{x}_{l+1,T} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{x}_{I,1} & \mathbf{x}_{I,2} & \cdots & \mathbf{x}_{I,T} \end{bmatrix} \quad (7)$$

ここで、 $l = r/100 \times I$ である。

これらを用いて、検証で用いる NN モデルの場合のクロスバリデーションのフローを図-3 に示した。

- A. 学習に用いるハイパパラメータを決定する。過学習を抑制するのであれば、モデルの実質的な自由度を下げれば良いから、ニューロン数を削減するか、ドロップアウト率を上げる。そのように過学習を抑制し過ぎると、今度は予測精度が低下する(ハイバイアス)トレードオフの関係にあるため、過学習せず、予測精度が最大のハイパパラメータを探索する。
- B. まず $\mathbf{D}_L$ を用いて結合荷重パラメータ行列 $\mathbf{w}_{D_L}$ を推定(学習)する。学習済みのモデルによる、学習データの予測関数は $\mathbf{y}(\mathbf{D}_L, \mathbf{w}_{D_L})$ と書く事が出来る。

図-4 過学習とハイバイアスの検知

- C. 学習済みの結合荷重パラメータ $\mathbf{w}_{D_L}$ を用いて、RNN モデルにより試験データの予測を出力する。予測関数は、 $\mathbf{y}(\mathbf{D}_T, \mathbf{w}_{D_L})$ である。
- D. 予測精度を表す指標として学習データと、試験データで別々に、次の 2 乗平均平方根誤差(RMSE: Root Mean Square Error)を求める。試験データの場合、RMSE は次式から求められる。

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{m=1}^M (y_m - d_m)^2} \quad (8)$$

ここで、 $y_m (m=1,2,\dots,M)$ は $\mathbf{y}(\mathbf{D}_T, \mathbf{w}_{D_L})$ の予測値の 1 つ、 d_m は y_m に対応した点検結果の実測値である。この RMSE は誤差を表すため、値が小さいほど予測精度は高い事を示す。

- E. RMSE の結果から、過学習及び予測誤差のハイバイアスは図-4 のように判断できる。学習繰り返しの回数が増加するほど学習データの RMSE は低下する一方、試験データの RMSE が増加する場合、過学習により予測性能が低下していると判断する。過学習は発生していないが、様々なハイパパラメータの下で学習を行った試験データの RMSE が、相対的に高ければ、ハイバイアス状態だと判断する。もし過学習又はハイバイアスであれば、A. からもう一度処理を行う。それらの問題が発生していなければ、そのハイパパラメータを採用する。

これにより決定したモデルの、試験データに対する RMSE をモデルの予測精度とする。

なお、実際は学習データと試験データは、それぞれのデータ内で学習毎にランダムに並び替えるが、本稿では説明の簡単の為省略した。

c) 提案するクロスバリデーション手法

本研究で提案するクロスバリデーション手法は、前項の従来手法が $\mathbf{D}$ を行方向(単位区間方向)に分割して $\mathbf{D}_L$ 及び $\mathbf{D}_T$ を作成した事に対して、次のように列方向(時刻方向)に分割し、同様の評価を行うタイムスライスク

ロスバリデーションと称す。

$$\mathbf{D} = [\mathbf{D}_p \quad \mathbf{D}_F] = \begin{bmatrix} \mathbf{x}_{1,1} & \cdots & \mathbf{x}_{1,s} & \mathbf{x}_{1,s+1} & \cdots & \mathbf{x}_{1,T} \\ \mathbf{x}_{2,1} & \cdots & \mathbf{x}_{2,s} & \mathbf{x}_{2,s+1} & \cdots & \mathbf{x}_{2,T} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{x}_{I,1} & \cdots & \mathbf{x}_{I,s} & \mathbf{x}_{I,s+1} & \cdots & \mathbf{x}_{I,T} \end{bmatrix} \quad (9)$$

ここで、 s は任意の時刻であり、用いるデータにより決定する。 s を仮想的に現在と考えると、 $t \leq s$ の $\mathbf{D}_p$ は過去の点検結果データと考える事が出来る。従って、モデルは $\mathbf{D}_p$ により学習を行い、精度の評価は仮想的な将来のデータ($t > s$)である試験データ $\mathbf{D}_F$ により行う事で、実運用時に近い予測精度の評価が可能である。

d) 提案手法の必要性

提案手法の必要性は次の通りである。

- A. 点検データは全てのデータで共通する時刻情報を持つ。予測モデルの実運用時には、将来時刻の観測不可能な情報は完全に未知の状態であり、将来を予測する。提案手法で仮想的に過去のデータ $\mathbf{D}_p$ と将来のデータ $\mathbf{D}_F$ に分ける事で、より実運用に近い現実的な予測精度の評価を行う事が可能である。
- B. 仮想的な将来の時刻のデータを学習に用いる事が不適である理由は次の通りである。舗装の点検結果は、点検装置の変更、装置のキャリブレーションの状態、点検担当者の個人差等により系統誤差が変化する。また舗装の劣化自体が舗装の素材、交通量の変化、施工基準等により変遷する。将来を予測する上で、その様な観測不可能な変数の変化は現時点で未知のはずである。それらの情報を含む将来のデータを学習すれば、予測精度は当然良くなるが、実運用上は不可能であるため、予測精度を過大評価する事がある。
- C. RNN モデルのクロスバリデーションは、前項で述べた通り学習データと試験データの RMSE よりハイパパラメータを調整し、過学習とバイアスのバランスが取れるようにする。将来予測に適用する上で、様々な未知の変化を含む将来の状態に対して最適なハイパパラメータを決定する必要がある。本提案手法のデメリットは、予測したい時刻以上の時系列数が必要である点である。例えば1年間隔で点検されており、5年後までの予測精度を評価したい場合、 $s=(T-5)$ に設定する必要があるため、少なくとも5年以上の点検時系列データが必要である。

4. 提案手法の検証

(1) 川崎市の路面性状調査履歴データ

a) データの概要

表-1 入力データ項目

説明変数	前処理
補修後の経過年数	正規化
大型車交通量	
大型車累積交通量	
車道幅員	
予測対象区間の平均わだち掘れ量	
前後KPの平均わだち掘れ量	
予測対象区間の最大わだち掘れ量	
前後KPの最大わだち掘れ量	
予測年数	
舗装厚	
CBR(路床の支持力比)	ダミー変数化
前回調査から補修の有無	
交差点の有無	
構造物(橋又はトンネル等)	
最新の補修工法	

本研究で用いたデータは川崎市の1989年～2015年の路面性状調査履歴である。そのデータの中でモデルの入力(説明変数)として用いた項目は表-1に示した。予測、精度検証の対象となるモデルの出力は、単位区間の平均わだち掘れ量である。路面性状調査は、土木研究センターによって行われる路面性状自動測定装置の性能確認試験に合格した装置によって行われることで、点検結果が規定の精度範囲内に収まる事が保証されている。今回用いるわだち掘れ量の精度規定は、プロフィールメータによる実測に対して $\pm 3\text{mm}$ 以内である。このデータは、路線を50mごとの単位区間に分割し、単位区間当たり1件の点検結果データが記録されており、調査は殆どの単位区間において3年間隔で行われている。予測を行う際には、ある時刻の実測値を元に、予測年数に指定された年数後の予測値を出力する。RNNは過去のデータを全て用いて現時点の予測を行うため、予測する対象のその時刻までの履歴を全て用いて、予測年数後(3～9年後)の予測を行った。

比例尺度のデータは正規化を行い、名義尺度はダミー変数化した。補修後の経過年数は最新の補修記録からの経過年数を表し、補修記録が一度もないデータについては欠測値として学習した。大型車交通量は、その時点で最新の交通センサスの24時間累計交通量を用い、その値を年について累積したものが大型車累積交通量である。予測対象区間の平均わだち掘れ量は、その単位区間の平均わだち掘れ量を表し、予測対象区間の最大わだち掘れ量は、その区間の最大のわだち掘れ量を表す。前後KPの平均わだち掘れ量及び最大わだち掘れ量は、予測対象区間の前後区間両方の平均わだち掘れ量及び最大わだち掘れ量である。最新補修工法は最後に行われた補修工法(新設、打ち替え、切削カバー、オーバーレイ、路上再生の何れか)を示す。

b) 外れ値の除去

図-5 従来手法と比較するためのデータ分割

表-2 モデルのハイパパラメータ

項目	ModelC	ModelP
学習率	0.002	0.002
学習繰り返し回数 (epoch)	100	100
ミニバッチサイズ	50	50
ドロップアウト率	0.5	0.5
中間層数	2層	2層
中間層の ニューロン数	30	7

補修された箇所とされていない箇所では当然分布が異なるため、補修の記録がある箇所と無い箇所を分け、それぞれで平均わだち掘れ量の年あたりの変化量の上下5パーセンタイルを除去した。処理後の学習・試験・検証に用いた点検結果データ数は47,553件である。外れ値除去後のわだち掘れ量データの平均値は6.39mm、標準偏差は3.20mm、25パーセンタイルは4mm、50パーセンタイルは6mm、75パーセンタイルは8mmである。なお、記録最小単位が1mmであるため、1mm単位のパーセンタイル値となる。

(2) 従来手法との比較による検証

a) 従来手法と比較するためのデータ分割

舗装の設計寿命は通常10年であり、本研究で用いた殆どの点検時系列データのサンプリング間隔は3年である。NNは連続関数であるため、実測値が存在しない時刻についても、学習している9年後までの連続値を予測年数に与える事により、近似曲線として予測できる。重原ら⁷⁾はその性質を利用して限られたデータから大型車交通量に対するパフォーマンスカーブの近似曲線を求めている。しかし、点検が行われておらず実測値が無い時刻では予測精度の評価が不可能であるため、本研究では

実測値がある3時系列分、9年後までの予測精度の評価を行った。従って最新の調査結果が2015年であることから、 s は9年前の $s=2006$ とした。

本研究のように1万件以上の十分なデータ数がある場合、学習データと試験データの分割は、予測精度のばらつきを抑え、かつ学習データも出来るだけ多く確保するために、試験データを1~3割、残りを学習データとする事が多い。本来は時系列方向に試験データ、学習データの2つに分割する前述の提案手法が良いが、ここでは一般的なクロスバリデーション(従来手法)と比較するため、図-5に示すようにデータを4分割した。

学習データ D_{PL} は、 $t \leq 2006$ の条件下でランダムサンプリングされた80%とした。従来手法と提案手法は全く同じ D_{PL} を学習する。従来手法と提案手法の違いは試験データの違いである。従来手法は試験データとして、 $t \leq 2006$ の、学習データに使用しない残りの20%、 D_{PT} を用いる。従って、 D_{PT} に対するRMSEが従来手法による予測精度評価である。一方、提案手法は試験データとして、 $t > 2006$ のデータを50%ランダムサンプリングした D_{FT} を用いるため、 D_{FT} に対するRMSEが提案手法による予測精度評価になる。それらとは別に、学習・試験時にも用いない独立した検証データ D_V として、 $t > 2006$ の残りの50%を用いた。

RNNモデルは前述の通りクロスバリデーションによりハイパパラメータが変わるため、従来手法の試験データ D_{PT} を用いて構築したモデルをModelC、提案手法の試験データ D_{FT} を用いて構築したモデルをModelPとした。

b) 従来手法と提案手法の比較

ModelC、ModelPのハイパパラメータを表-2に、それらのモデルによる D_{PT} 、 D_{FT} 、 D_V を予測した結果から算出したRMSEを、予測年数ごとに図-6~8に示した。学習の効率化の為に、学習率を自動的に調整する効果があるADAM¹²⁾を用いた。簡単のため中間層のニューロン数は2層の中間層で同一である。

なお、精度評価に用いるRMSEは、実測値と真値との差である測定誤差を含む。測定誤差の主因の1つとして、測定装置の違いによる機差がある。機差は、2つの測定値間の差を平均した偏り(Bias)と、その標準偏差、ばらつき(Precision)に分けて考える事が出来る。ばらつきは予測モデルや平均化により減少させることが出来るが、偏りは減少させることが困難であるため、ここでは偏りについて考察する。阿部ら¹³⁾は性能確認試験に合格した2つのわだち掘れ量測定装置間の偏りが2.3mmである事を報告している。実際の測定値は多種の装置が混在している事、同一の装置であってもキャリブレーション状態によって変化する事から、より様々な偏りが生じる事を指摘している。また、わだち掘れ測定装置を保有する

TxDOT(Texas Department of Transportation)他 6つの機関による測定結果と、6ft ストレートエッジ法によるわだち掘れ量実測値(GT : Ground Truth)間の機差について検証した FHWA(Federal Highway Administration)の調査報告¹⁴⁾では、最も GT に近いとされる TxDOT(ASTM), Dynatest, Roadware の3機関の GT に対する偏りの平均値は 1.42mm である。阿部らにより検証が行われた 1993 年当時は、アナログフィルムにわだち掘れ形状を投影する方式であったが、FHWA の調査では、2000 年代後半に登場してきた INO 社の LRMS など、1 点の測定精度が±1mm 以内でより自動化された測定装置を用いた結果、わだち掘れ量の偏りが 1.42mm に抑えられたと考えられる。このことから、ModelP において図-6 の D_{PT} よりも図-7 の D_{FT} に対する RMSE の絶対値が小さくなっている要因は、新しいデータほどより高精度の装置を用いているためだと考えられる。

ModelP では、試験データである D_{FT} に対する過学習を抑えるため、ハイパパラメータのニューロン数は ModelC と比較して 1/3 以下になった。これは、仮想的な現在時刻を $t=s$ とすると、 $t \leq s$ の情報から $t > s$ の予測をする事は、 $t \leq s$ を予測するよりも難しい事を反映している。この事から、RNN モデルにおいては将来予測のためのハイパパラメータを探索する必要がある、3(2)(d)C.の有効性が確認できた。

ModelC においては、実際に将来予測を行った際の精度である D_{FT} 又は D_V に対する RMSE より、 D_{PT} に対する RMSE の方が小さいため、予測精度を過大評価する事になる。ModelC が予測精度を過大評価している量は、各予測年数ごとに(D_{PT} に対する RMSE) - (D_{FT} に対する RMSE)を求める事により、本検証においては 0.26～0.33mm の範囲である。このことから、3(d)B.で指摘した従来手法では予測精度を過大評価する場合がある事について確認できた。

また、それらの事から提案手法による予測精度評価に当る D_{FT} の RMSE を評価した図-7 では、従来手法より実運用に近い、現実的な将来の予測精度の推定が出来たと考えられるため、3(2)(d)A.が検証できた。他方で、将来を予測するのではなく、例えば近隣の路線の状態を予測する等、学習データと同じ時刻内の予測を行う場合は、従来手法を用いてハイパパラメータの決定を行った方が予測精度は高いと言える。

ほかにも、従来手法と提案手法の 6 年後と 9 年後の予測と比較し、3 年後の RMSE の差は大幅に小さい。これは、3 年後の RMSE が 6 年後、9 年後と比較し明らかに小さい事から解るように、予測が容易であるため、RMSE に含まれる測定誤差に対して予測誤差が相対的に小さく、差が出難いためだと考えられる。

さらに、図-7 の D_{FT} に対する RMSE と、図-8 の D_V に

図-6 D_{PT} に対する予測年数ごとの RMSE

図-7 D_{FT} に対する予測年数ごとの RMSE

図-8 D_V に対する予測年数ごとの RMSE

対する RMSE は殆ど差がないため、実用上は本章のように試験データと検証データを別に用意する必要は無く、学習データと試験データに時刻方向で2分割する提案手法で十分である事が解った。

5. まとめ

本研究は、時刻方向にデータを分割する事による、舗装の点検時系列データを用いた劣化予測に対する、より現実的な予測精度の評価方法を提案したものである。本研究で得られた知見は以下の通りである。

- A. 提案手法では、実運用時の劣化予測で行われるように仮想的な過去のデータから将来のデータを予測して RMSE を求める事で、より実運用の劣化予測に近い評価が可能である事が検証できた。
- B. 従来のクロスバリデーション手法では、データを

ランダムに分割する事で、学習データに含まれる時刻のデータで精度評価を行うため、予測精度を過大評価する場合がある事が解った。

- C. 提案手法でハイパパラメータの調整を行う場合、現実的な将来予測の難しさから、従来手法よりも大幅に過学習を抑制する必要がある。RNN モデルでは、ニューロン数を 1/3 以下にする必要があった。

謝辞：本研究を遂行するにあたり、川崎市の貴重な路面性状調査履歴データを使用させて頂いた。ここに記して謝意を表する。

参考文献

- 1) 谷本圭志, 杉本泰亮, 宮本慎也, 灘英樹, 細井由彦: 集約的維持管理方策に関するモデル分析—道路舗装と下水道管路を対象として—, 土木学会論文集 G (環境), Vol.67, No.6, pp.83-91, 2011.
- 2) 小川泰弘, 上田孝行, 森地茂: 道路舗装の維持管理基準の決定に関する研究, 土木計画学研究・講演集 (CD-ROM), Vol. 39, 2009.
- 3) 宮崎文平, 加藤寛之, 小濱健吾, 貝戸清之, 風戸崇之, 田中克則: 膨大な路面性状調査データに基づく舗装補修施策の評価法の提案, 土木学会論文集 F4 (建設マネジメント), Vol. 71, No. 3, pp. 142-161, 2015.
- 4) Abaza, K. A. : Back-calculation of transition probabilities for Markovian-based pavement performance prediction models, *International Journal of Pavement Engineering*, Vol. 17, No. 3, pp. 1-12, 2014.
- 5) 石谷雅彦, 中川伸一, 小笠原章, 高橋守人, 笠原篤: 美々新試験道路におけるアスファルト舗装のパフォーマンスに関する研究, 土木学会論文集, No.564, pp.265-276, 1997.
- 6) 斎藤雅道, 福田正: 少数データによる舗装のパフォーマンスモデルの構築, 舗装工学論文集, Vol. 1, pp. 181-186, 1996.
- 7) 重原大二郎, 西澤辰男, 中源達雄, 平野誠志: ニューラルネットワークによる北陸地方のアスファルト舗装わだち掘れ進行モデル, 土木学会舗装工学論文集, Vol.13, pp.25-30, 2008.
- 8) Tabatabaee, N., Ziyadi, M., and Shafahi, Y. : Two-Stage Support Vector Classifier and Recurrent Neural Network Predictor for Pavement Performance Modeling, *Journal of Infrastructure Systems*, Vol.19, No.3, pp.266-274, 2013.
- 9) 奥田知之, 鈴木康豊, 神武直彦: リカレントニューラルネットワークモデルによる路面性状値予測の検討, 土木情報学シンポジウム, 講演集, Vol.41, pp.159-161, 2016.
- 10) Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. : Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling, *arXiv*, pp.1-9, 2014.
- 11) Semeniuta, S., Severyn, A., and Barth, E. : Recurrent Dropout without Memory Loss, *arXiv*, pp.1-11, 2016.
- 12) Kingma, D. P., and Ba, J. L. : Adam: a Method for Stochastic Optimization., *International Conference on Learning Representations*, 2015.
- 13) 阿部頼政, 飯野忠雄: わだち掘れ測定データの解析法に関する研究, 土木学会論文集, 公益社団法人 土木学会, No. 478, pp. 117-123, 1993.
- 14) Serigos, P. A., Prozzi, J. A., Nam, B. H., and Murphy, M. R. : *Field Evaluation of Automated Rutting Measuring Equipment*, Center for Transportation Research, pp. 112-154, 2012.

PROPOSAL AND VERIFICATION OF ACCURACY EVALUATION METHOD OF PAVEMENT DEGRADATION PREDICTION MODEL

Tomoyuki OKUDA, Kouyu SUZUKI and Naohiko KOHTAKE

In order to realize effective pavement maintenance and management, it is necessary to predict the future state change by prediction of pavement deterioration and to optimize the maintenance and management cost such as inspection, repair, renewal etc. Various pavement deterioration prediction models have been proposed, but it is not easy to determine under what circumstances each model is best.

Therefore, this research proposes time slice cross validation that can compare realistic prediction accuracy of prediction model to inspection history data having time information with each other. The proposed method divides time series inspection data in the time direction, estimates parameters of the model from the learning data, and performs realistic prediction accuracy evaluation by using the prediction result of the test data using the parameters. We applied the proposed method to pavement deterioration prediction by RNN model and verified that future prediction accuracy can be more realistically evaluated than conventional method.