

次元圧縮を適用した リアルタイム浸水予測モデルの複数地点への適用

Application of real-time inundation prediction model
with dimensional compression to multiple locations

修士 中山龍也¹⁾羽物裕人²⁾一言正之³⁾樺山和男⁴⁾
Ryuya Nakayama, Yuto Habutsu, Masayuki Hitokoto and Kazuo Kashiyama

¹⁾中央大学大学院 理工学研究科 (〒 112-8551 東京都文京区春日 1-13-27, E-mail: a20.3emg@g.chuo-u.ac.jp)

²⁾工修 中央大学大学院 理工学研究科 (〒 112-8551 東京都文京区春日 1-13-27, E-mail: a19.6t5r@g.chuo-u.ac.jp)

³⁾工博 日本工営株式会社 河川水資源事業部 河川計画部 (〒 102-8539 東京都千代田麹町 5-4, E-mail: hitokoto-ms@n-koei.jp)

⁴⁾工博 中央大学 教授 (〒 112-8551 東京都文京区春日 1-13-27, E-mail: kaz@civil.chuo-u.ac.jp)

This paper presents a real-time flood prediction model using dimension reduction methods with Deep Neural Network (DNN) that can handle overflow from multiple points. Three-dimension reduction methods, Singular Value Decomposition (SVD), Non-Negative Matrix Factorization (NMF), and Auto Encoder (AE), are employed and evaluated in terms of computational efficiency and prediction accuracy using a flood simulation of the Arakawa River. The results demonstrate that the proposed method not only significantly reduces computation time but also maintains high prediction accuracy.

Key Words : Deep Learning, Flood Prediction, Dimensional Compression, Simulation

1. はじめに

近年、豪雨による河川の氾濫が頻発しており、今後は地球温暖化によって氾濫リスクが増大することが懸念される。洪水氾濫による被害を軽減するには、流域の降雨量や流量情報を迅速に把握し、適切な洪水対策を講じることが重要である。さらに、リアルタイム予報の観点から計算時間の課題を改善するため、数値解析手法の代理モデルとして機械学習を用いた浸水予測の研究が進められている。

I-Feng ら [1] は、実際の浸水域データを Stacked AutoEncoder (SAE) により低次元の特徴量へ圧縮し、再帰型ニューラルネットワークと復元器を組み合わせて降雨情報から浸水域を予測するモデルを構築した。一方、島袋ら [2] は降雨流出氾濫モデルの計算結果を特異値分解で圧縮し、深層学習 (DNN) と復元器を用いて降雨情報から浸水域を予測する手法を提案している。また、著者ら [3] は次元圧縮手法を用いたリアルタイム浸水域予測モデルを構築し、3 種類の次元圧縮手法の有効性を検討してきた。しかし、既往研究では予測の適用範囲が 1 地点からの越流に限定されていた。

そこで本論文では、複数地点からの越流を考慮し、さらに異なる 3 つの次元圧縮手法を適用することでケーススタディを実施し、各条件における予測精度と計算時間を比較して、浸水域予測モデルの条件設定に応じた性能を評価した。

2. 学習データについて

モデル構築のフローチャートを図-1 に示す。

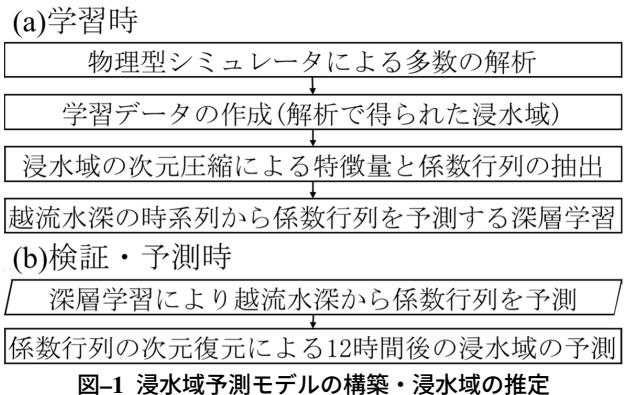


図-1 浸水域予測モデルの構築・浸水域の推定

(1) モデル構築と解析の手順

浸水域予測モデルは、次元圧縮手法と深層学習に基づく予測器を組み合わせることで構築した。図-1 中の (a) では学習時に以下の手順でモデルの構築を行った。

- 越流箇所や浸水規模など、様々な浸水状況シナリオに基づいた物理法則に沿った洪水氾濫解析を実行。
- 洪水氾濫解析の結果をもとに、浸水域の学習データセットを作成。
- 次元圧縮手法を用いて、浸水域の学習データセットから特徴量と係数行列を抽出。
- 1 時間ごとの越流水深の時系列データから係数行列を推定する深層学習モデルを構築。

以上の手順により構築した浸水域予測モデルを用いて、浸水範囲を推定する。本研究は、CPU : Core i7-12400F, GPU : RTX 3070 を使用し、Python 3.9 を基盤に TensorFlow, NumPy, scikit-learn で実装した。

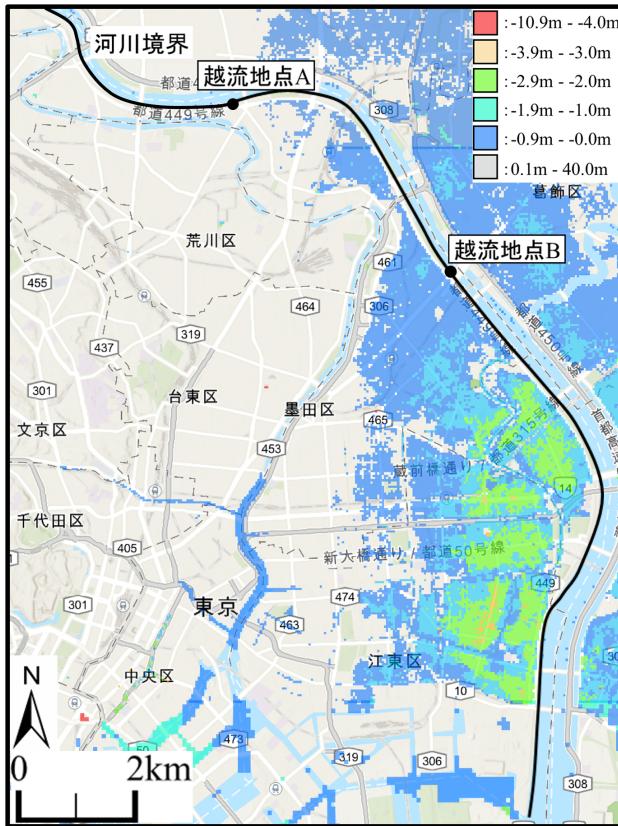


図-2 対象領域とその標高

(2) 浸水解析モデル

氾濫浸水解析には、氾濫水の挙動を高精度に表現できる Dynamic Wave 法を適用した。[4] 浸水解析モデルの x , y 方向の運動量保存式および質量保存式は以下のとおりである。

$$\frac{\partial \mathbf{M}}{\partial t} + \frac{\partial(\mathbf{uM})}{\partial x} + \frac{\partial(\mathbf{vN})}{\partial y} + gh \frac{\partial \mathbf{H}}{\partial x} = -gh \frac{n^2 u \sqrt{u^2 + v^2}}{R^{\frac{4}{3}}} \quad (1)$$

$$\frac{\partial \mathbf{N}}{\partial t} + \frac{\partial(\mathbf{uN})}{\partial x} + \frac{\partial(\mathbf{vN})}{\partial y} + gh \frac{\partial \mathbf{H}}{\partial y} = -gh \frac{n^2 v \sqrt{u^2 + v^2}}{R^{\frac{4}{3}}} \quad (2)$$

$$\frac{\partial \mathbf{h}}{\partial t} + \frac{\partial \mathbf{M}}{\partial x} + \frac{\partial \mathbf{N}}{\partial y} = 0 \quad (3)$$

ここで、 $M = uh$ は x 方向の単位幅流量、 $N = vh$ は y 方向の単位幅流量、 u は x 方向の流速、 v は y 方向の流速、 h は水深、 H は基準面からの水位、 R は径深、 g は重力加速度、 n はマニングの粗度係数である。なお、空間方向は有限差分法、時間方向は陽解法を用い、時間増分量 25m メッシュで 0.2 秒、粗度係数は基盤地図情報の土地利用分類 [5] に基づいて設定した。

(3) 泛濫解析条件の設定

本研究では、2 地点からの越流における 1 時間ごとの越流水深を入力データとし、洪水氾濫解析で得られた浸水深データを学習データとした。図-3 は、平均越流水深が最大の場合と最小の場合の時系列変化を示し、図-4 は氾濫解析による結果を示す。なお、越流水深の

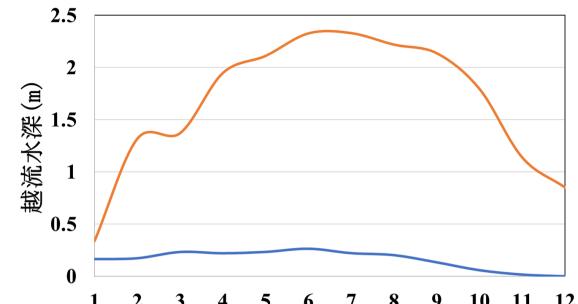


図-3 越流水深データについて

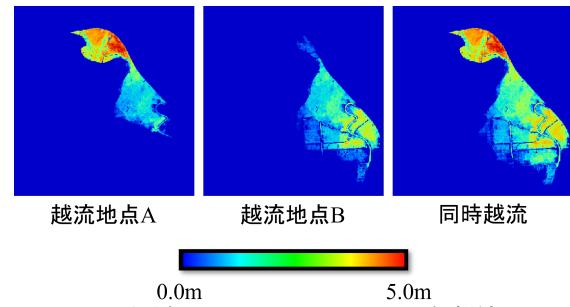


図-4 泛濫シミュレーションによる解析結果

時系列データは正弦波にノイズを付与したものであり、以下の式 (4) により生成した。

$$h = \left\{ \sin \left(a - 0.5 + \frac{\pi t}{720} \right) + 0.3b \right\} \times 2c \quad (4)$$

ただし、式 (4) の計算において $h < 0$ の場合は 0 に置換した。ここで、 a , b , c はそれぞれ 0~1 の乱数、 t は時間 (分) を表す。対象地域は荒川下流デルタ地帯とし、越流地点および標高分布を図-2 に示す。標高・土地利用データは国土地理院の国土数値情報 [5] を用い、25m メッシュにおいて最近隣内挿法を適用して取得した。解析条件は、地点 A および B それぞれからの越流、ならびに両地点からの同時越流の 3 条件とし、各条件で 1000 ケースの氾濫解析を実施した。図-4 は、各解析条件で越流水深が最大となるケースの結果を示す。なお、氾濫解析の計算時間は 1 ケースあたり最大 213 秒であった。

(4) 次元圧縮

次元圧縮とは、高次元のデータの重要な特徴を保持しつつ低次元で表現する手法である。データの本質的特徴を維持しながら次元数を削減することで、深層学習の推論に要する計算効率が向上する。本研究では、シミュレーションの格子点数を圧縮前の次元数とみなし、特異値分解 (以後 SVD)[6]、非負値行列因子展開 (以後 NMF)[7]、AutoEncoder (以後 AE)[8] を浸水域データに適用した。表-1 および表-2 に各次元圧縮手法の実行条件および AE モデルの学習条件を示す。

(5) 深層学習手法

本研究では、DNN (Deep Neural Network) を用い、1 時間ごとの越流水深を説明変数、各ケースごとの係数行

表-1 次元圧縮の実行条件	
対象データ	3000 ケースの氾濫域解析結果
圧縮前の次元数	221392 次元 (404×548)
次元圧縮手法	特異値分解 非負値行列因子展開 AutoEncoder
圧縮後の次元数	1 から 20 次元 (1 次元刻み)
評価指標	RMSE

表-2 オートエンコーダモデルの学習条件	
訓練データ	2400 ケース
テストデータ	600 ケース
学習回数	50 回
中間層の数	2 層 (256, 128node)
バッチサイズ	32
検証方法	交差検証

表-3 DNN モデルの学習条件	
訓練データ	2400 ケース
テストデータ	600 ケース
説明変数	越流水深
目的変数	係数行列
学習回数	最大 500 回 (損失が 10 回で 低減しない場合に早期終了)
中間層の数	4 層 (128, 64, 32, 16node)
バッチサイズ	32
損失関数	平均二乗誤差
検証方法	交差検証

列を目的変数とした深層学習モデルを構築した。表-3 に DNN モデルの学習条件を示す。

(6) 浸水域予測の評価指標

浸水域予測の評価指標として、二乗平均平方根誤差 (Root Mean Square Error ; RMSE) を式 (5) に示す。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (a_i - f_i)^2} \quad (5)$$

ただし、 n は格子点数、 a_i は予測した浸水深、 f_i はシミュレーションによる浸水深である。

3. ケーススタディ結果

各次元圧縮手法を適用した浸水域予測モデルの予測精度および予測時間について比較を行った結果は、以下のとおりである。

(1) 予測精度について

図-5 に、圧縮次元数と RMSE の関係を示す。図中の点線は DNN を介さず各次元圧縮手法を用いて浸水域データセットを圧縮し、その後復元した結果を示している。一方、実線は DNN により予測された係数行列を

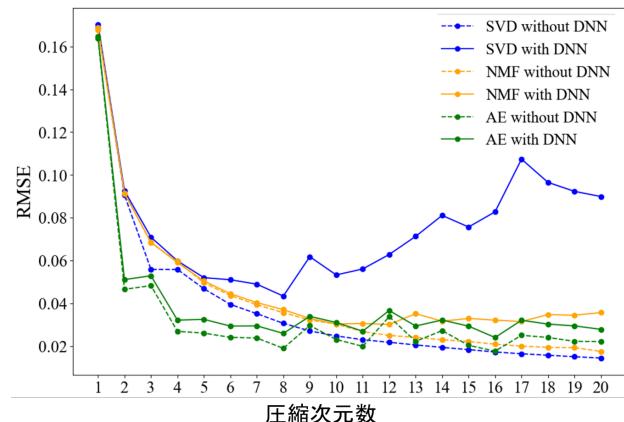


図-5 RMSE と圧縮次元数の関係について

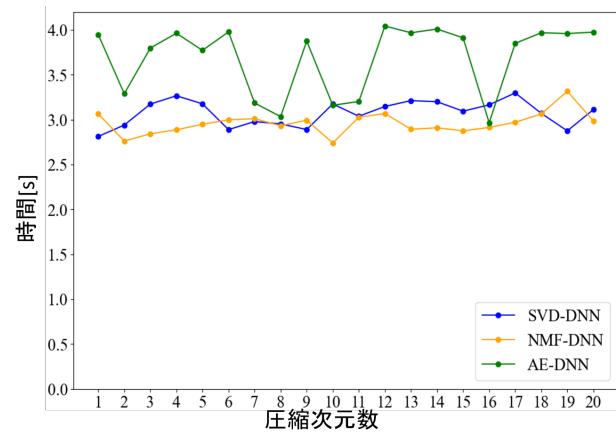


図-6 圧縮次元数と予測時間の関係

用いて浸水域を復元した結果である。図-5 から、DNN を介さずに復元する場合、いずれの手法においても圧縮次元数を増加させることで RMSE の値が減少する傾向があることが確認できる。これは、圧縮次元数を増加させることで浸水域の特徴をより柔軟に表現できるためである。一方、DNN を介して予測された係数行列を用いた場合、いずれの次元圧縮手法を適用した予測モデルも、圧縮次元数をある程度増加させると RMSE が低減するものの、過度に増加させると性能が悪化する傾向も見られた。本研究では、10 次元程度が最適な圧縮次元数であると判断した。また、浸水域予測において DNN を用いた場合と DNN を用いない場合で RMSE の値に差が生じたのは、DNN による予測誤差が影響していると考えられる。

(2) 予測時間に関する検討

図-6 に、浸水域の予測時間と圧縮次元数の関係を示す。予測時間は、DNN による推論と次元復元に要した時間の合計である。SVD および NMF をそれぞれ適用したモデルでは予測時間が約 3 秒、AE を適用したモデルでは約 4 秒と、いずれも元の氾濫解析と比較して短時間での予測が可能である。図-6 から、圧縮次元数の変化が予測時間に大きな影響を及ぼさないことが確認された。これは、元データが 221392 次元という極めて大きな次元数であるため、10 次元、100 次元、1000 次

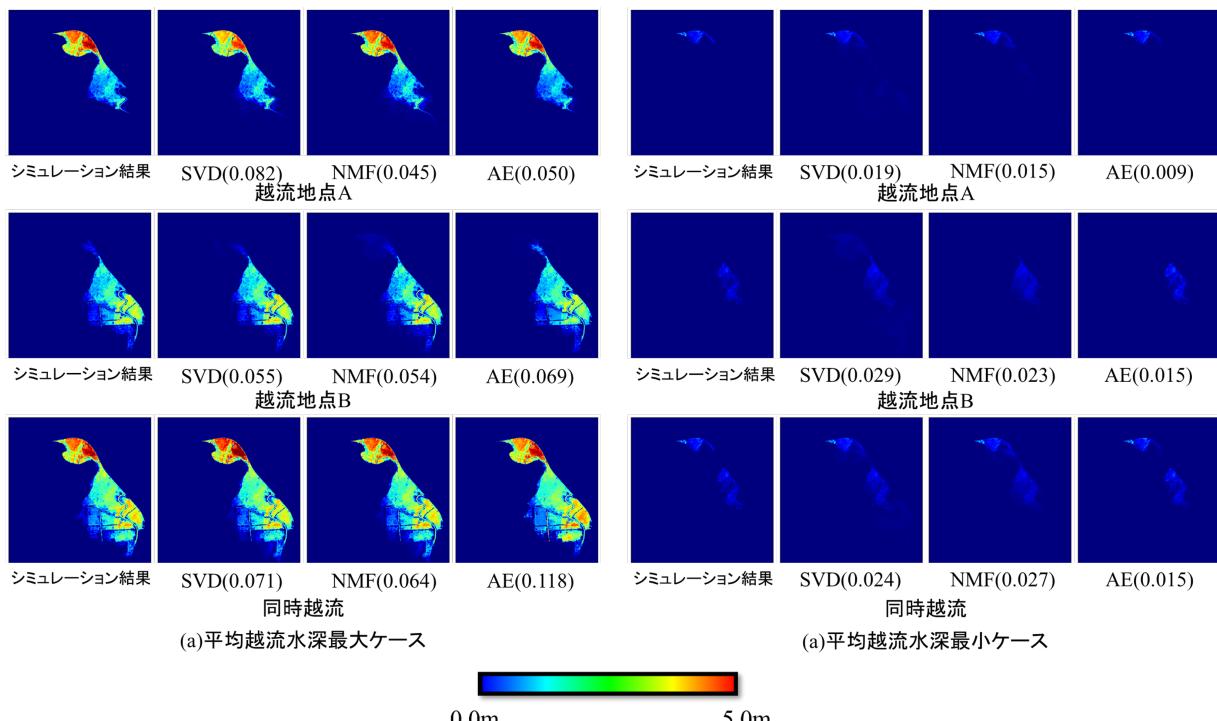


図-7 各次元圧縮手法を適用した氾濫域の予測結果 (10 次元)

元と次元を拡大すれば予測時間は確実に増加すると考える。しかし、本研究では予測時間と精度の両立を考慮し、1~20次元の範囲内で検討した結果、22万次元と比較すると次元数の差はごく僅かであり、予測時間に顕著な違いが現れなかったと考える。

(3) 10次元における各条件での予測結果

図-7(a)は、平均越流水深が最大の場合に3つの異なる次元圧縮手法を適用した浸水域予測モデルの予測結果を示し、図-7(b)は平均越流水深が最小の場合における予測結果を示す。各条件の図下には、用いた次元圧縮手法とRMSEの値が記載されている。図-7(a)および(b)から、越流水深が最大の場合、NMFを適用したモデルがすべての条件下で最も低いRMSEを示し、越流水深が最小の場合はAEを適用したモデルが最小のRMSEを示した。一方、SVDは他の手法と比較するとやや精度が劣るもの、大域的な浸水形状の予測には有用であると考えられる。

4. おわりに

本論文では、2地点からの越流を考慮した次元圧縮を適用した浸水域予測モデルを構築し、その性能評価を行った。3つの異なる条件の洪水氾濫解析で得られた浸水深分布に対し、3種類の次元圧縮手法を適用した予測結果から、以下の結論が得られた。

- 2つの氾濫開始点における浸水域データに対して最も、越流水深が最大・最小の場合において浸水域予測は可能であることを確認した。
- DNNに次元圧縮手法を適用することで、浸水域の予測時間を大幅に短縮しながら精度を維持できる。
- 2地点からの越流においては、およそ10次元程度

で十分な浸水域予測が可能である。

今後は、他の河川への適用や越流地点の増加による検証を通じ、次元圧縮を適用した浸水域予測モデルの汎化性能のさらなる向上を目指す。

参考文献

- [1] I-Feng K, Yanlai Z, Li-Chiu C, Fi-John C.: Exploring a Long Short-Term Memory based Encoder-Decoder framework for multi-step-ahead flood forecasting, *Journal of Hydrology*, 2020.
- [2] 島袋隆也, 塩尻大也, 小楢峻司:深層学習モデルを用いた浸水深の時空間分布予測のエミュレーション, *AI・データサイエンス論文集* 4巻3号, p.553-560, 2023.
- [3] 中山龍也, 羽物裕人, 一言正之, 横山和男:深層学習を用いたリアルタイム浸水域予測における次元圧縮の適用, *AI・データサイエンス論文集* 5巻3号, p. 563-571, 2024.
- [4] 一言正之, 荒木光一, 古木宏和:敵対的生成ネットワークによる洪水氾濫浸水域の推定, *人工知能学会全国大会* 第33回, 2019.
- [5] 国 土 地 理 院, 国 土 基 盤 情 報 <https://www.gsi.go.jp/kiban/> (参照 2023-10-10)
- [6] Eckart C. and Young G., The approximation of one matrix by another of lower rank, *Psychometrika*, 1(3), 211-218, 1936.
- [7] Lee D. D. and H. S. Seung, Learning the parts of objects by non-negative matrix factorization, *Nature*, 401(6755), 788-791, 1999.
- [8] Rumelhart D. E., G. E. Hinton and R. J. Williams, Learning representations by back-propagating errors. *Nature*, 323(1), 533-536, 1986.