

## 気象データを用いた高潮予測のためのディープニューラルネットワークの構築に関する検討

九州大学 学生会員 ○清水功基

九州大学 児玉充由 井手喜彦

九州大学 フェロー 橋本典明

### 1. はじめに

温暖化が進むと海面水温の上昇により、現在に比べて強い台風の発生頻度が高くなることや海面が上昇することが指摘されている。そのため、高潮災害への対応の重要性は今後さらに高まっていくと考えられ、より迅速で正確な判断のための高潮予測が求められている。現在、高潮予測には数値モデルによるシミュレーションが主流となっている。しかし、数値シミュレーションでは計算精度と計算速度がトレードオフの関係にあり双方を両立させるのは難しい。そこで本研究では、数値シミュレーションに代わってディープニューラルネットワーク(DNN)を使用した高潮予測の可能性に注目した。DNN モデルは高精度な数値シミュレーションの結果を事前に学習させることでその数値シミュレーションの精度を維持したまま高速で高潮予測を行うことができる。この高速な計算を活かし台風経路の不確実性を考慮したアンサンブル実験を行い確率論的な高潮予測などが可能となる。したがって、本研究では気象予報データを用いた高潮予測のためのディープニューラルネットワークの構築に関する検討を行った。

### 2. 研究手法

#### 2-1. 使用データ

DNN モデルの構築には多くの訓練データが必要なため「地球温暖化対策に資するアンサンブル気候予測データベース(d4PDF)」を用いた。まず、経度 130~133 度、緯度 32~35 度で囲まれる領域(図 1 の小領域)を通る台風を d4PDF から抽出した。それらが経度 128~138 度、緯度 30~36.6 度で囲まれる領域(図 1 の大領域)に入った時から、出るもしくは消滅した時刻までの毎時の気圧、風分布を外力として数値シミュレーションを行い、高潮偏差を得た。なお、シミュレーションモデルには周防灘での高潮推算に十分な精度を有する園田<sup>1)</sup>によって用いられた Finite Volume Community Ocean Mode(FV-COM)を用いた。

#### 2-2. DNN モデルの概要

DNN モデルは入力層、出力層、およびその中間に配置される複数の隠れ層からなる回帰モデルを使用した。入力層には図 1 の赤破線で囲む領域(DNN 領域)の気象データ(気圧・風)を入力する(格子数 31(緯度方向)×54(経度方向))。ただし、各格子の毎時の気象データを全て入力すると入力ノード数が膨大となり学習に要する計算コストの観点から現実的でないのに加え、台風毎に入力ノードを変える必要があるため一つのモデルでは対応できない。したがって、気圧であれば、対象とする期間における各格子の最小値分布を入力データとして与えた。例えば、図 1 の黒線が示す台風経路であれば気圧の入力データは図 2 のようになり、台風の経路上で気圧が低くなっていることがわかる。風についても同様に東西および南北風速の最大値分布を作成し、入力データに与えた。しかし、これらの入力データだけでは高潮を考慮する際に重要となる台風の移動速度の情報が含まれていない。そのため、台風が各格子内へ滞在した時間の空間分布も入力データとして使用した。したがって、構築した DNN モデルの入力ノード数は 31(緯度方向格子数)×54(経度方向格子数)×4(入力データの種類)=6696 である。なお、出力層は 1 ノードであり、宇部の最大高潮偏差とした。

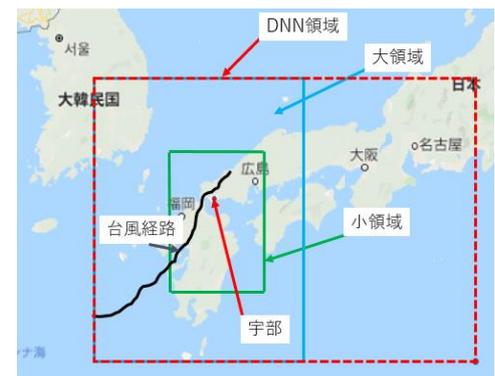


図 1 各領域とある台風経路

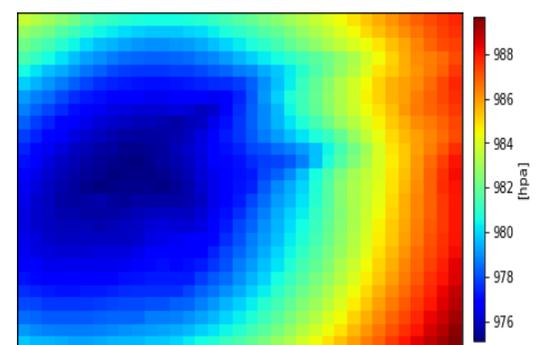


図 2 DNN 領域と台風経路の最小気圧分布

### 3.DNN モデルの最適化に関する検討

#### 3-1.適した DNN パラメータ

DNN パラメータである重み, 活性化関数および学習率の決定法を変化させたときの二乗平均平方根誤差(RMSE)を確認し, 適したパラメータを決定した. この検証のみ, DNN が学習を行っているか検証するため, 訓練データと検証データは同一のものを使用した. なお, DNN の中間層は2層とし各層 1000 ノードとした. 表1に活性化関数は ReLU, 学習率には adam を設定し, 重みの初期化法の違いによる誤差の比較を示す. gauss が最も精度が良かったため, この条件で活性化関数を変化させ精度を確認した(表2). その結果, Parametric ReLU の精度が最もよくなった. この設定を引き継いで学習率の決定法による精度の変化を調べ, 最も精度がよい Nesterov momentum を設定した(表3). 今回は計算コストの関係上, 上記の方法で局所的な最適パラメータを求めた. 以下の検討はここで決定した設定の DNN で実施した.

表1 重みの初期化の違いによる誤差の変化

重みの初期化	gauss	gaus_sqr(n)	lecun	gh_uniform	gb_gauss	hereh
RMSE(cm)	6.38	7.25	6.83	8.78	8.14	8.29

表2 重みの初期化の違いによる誤差の変

活性化関数	sigmoid	tanh	ReLU	Leaky ReLU	Parametric ReLU
RMSE(cm)	6.46	27.44	6.45	7.16	6.41

表3 学習率の違いによる誤差の変化

学習率	0.01	momentum	Nesterov momentum	adagrad	adadelta	rmsprop	adam
RMSE(cm)	3.19	2.08	0.92	3.88	2.52	5.92	3.66



図3 DNN 領域

#### 3-2.最適領域

DNN に使用する最適な領域を決定するため, 領域を変えた Case0~Case10 で精度を比較した. Case0 (図3の赤枠) は東経 128~138 度, 北緯 30.0~36.5 度. Case1~Case9 は Case n に対して東経  $128+0.3n \sim 138-0.7n$  度, 北緯  $30+0.4n \sim 36.6-0.1n$  度のように宇部を中心に領域を狭めていき, 最小領域の Case10 は東経 131~131.4 度, 北緯 33.8~34.0 度まで領域を絞った(図3). 図4に RMSE の領域依存性を示す. 最も精度がよいケースは Case3 である. Case3 より広域のケースでは情報が多いにも関わらず精度が落ちている. これは, 重みの数が増え更新頻度が減るためだと思われる. また, Case5 以降の精度が落ちたのは領域が狭いため予測に必要な情報が十分に含まれていないためと考えられる. これより, 本研究の最適な DNN 領域は Case3 であることが分かった.

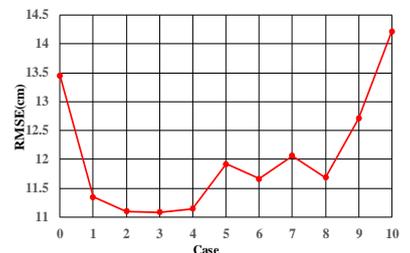


図4 RMSE の領域依存性

#### 3-3.入力データ数

入力データ数の増加と精度の依存性を確認するため, 検証データ数を 100 個に固定し訓練データ数を 50 から 450 個まで 50 個ずつ増加させた(図5). 訓練データ数が増加すると RMSE が減少することが分かる. ここでは, 訓練データ数 450 個まで検討したが, さらに訓練データ数を増やすことで精度の向上が予想される.

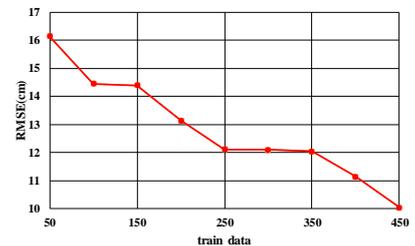


図5 RMSE の訓練データ数依存性

#### 3-4.入力データの種類

入力データの種類を減らすと DNN の学習計算コストを減らすことができる. そこで, 風と気圧分布は相関があるため, 一方の分布を使用しない場合を検討した. その結果, RMSE は気圧分布を使用しない場合は 11.35cm, 風速分布を使用しない場合は 13.21cm であり, どちらの分布も使用した場合の 9.92cm 比べ精度が落ちた. これより, 精度の高い予測には風と気圧分布の両方を入力データとして使用する必要があることが分かった.

### 4.終わりに

本研究では DNN モデルの最適なパラメータや領域の設定, またデータの種類や数について検討することで DNN モデルの精度を高めた. 今後, データ数の増加によりさらに精度を高めることができると思われる. 本研究で開発した DNN モデルの高速で高精度な高潮推定は, 確率論的な高潮予測の検討や高潮を引き起こす台風の最悪経路の推定に活用することが出来ると思う.

#### 参考文献

1) 園田ら(2019): 周防灘における将来の高潮に関する研究, 土木工学論文集 B2 (海岸工学), Vol.71, No.2, I\_1183-I\_1188.