アンサンブル気象予報データを用いた機械学習とベイズ推定による波浪予測システムの構築

九州大学	学生会員	○大橋果歩
九州大学		児玉充由
九州大学	正会員	井手喜彦
九州大学	フェロー	橋本典明

1. はじめに

波浪予測は港湾・海洋工事の作業実施の可否判断などに必要とされる.近年,気象庁は気象に関するアンサンブ ル数値予報(27 ケース)の高分解能日本域 GPV データの提供を開始した.これにより,気象データから複数の波 浪予測値を得ることが可能になった.そこで,本研究ではこれらのアンサンブル波浪予測データに対してベイズ推 定を適用することで,より高精度な1つの最適推定値を導くことを試みた.また,気象データから波浪を予測する 際,数値モデルを用いた予測では27 ケースの計算に高い計算コストを要する.そこで将来的には実務への利用も見 据え,近年波浪予測に適用されその有用性が確認されている¹⁰機械学習のディープニューラルネットワーク(以下, DNN とする)を用いて,現実的な計算コストかつ高精度な波浪予測システムの構築を目指した.

2. 波浪予測システムの概略

図1に構築した波浪予測システムの流れ を示す.予測値の導出は大きく2つのステ ップに分けられる.まずステップ1では, 週間アンサンブル数値予報モデルGPV(27 ケース)をそれぞれ入力値とするDNNを 構築し,27ケースの波浪予測値をそれぞれ 導出する.なお,週間アンサンブル数値予 報モデルGPVは11日間先までの気象予報 データであるが、本研究では6日後までの





データを使用した.次いでステップ2では、ステップ1で得られた複数の予測値からより精度の高い1つの予測値を 推定する.具体的には、構築した27ケースのDNN それぞれに対して精度の評価を行い、ベイズ推定の一種である 線形最小分散推定により、最も精度の高い1つの予測値を導く.

3. ディープニューラルネットワークモデル

3-1. モデルの概要

DNN モデルは,隠れ層を 2 層に設定し,入力層には予測したい地点にお ける週間アンサンブル数値予報モデル GPV (27 ケース)の風速・風向データ を与え,出力層では波高を出力する.教師データには同地点の波高データ(観 測値)を使用した.なお図 2 のように,ある時刻の波高を予測する際,予測 したい時刻から過去1日間の風速・風向の予測データを入力値に用いた.ま た,DNN において予測精度を大きく左右するハイパーパラメータ(隠れ層の ノード数,エポック数など計8つ)については,ベイズ最適化を用いた探索 を行い値を設定した.

3-2. 精度検証

本研究では、地域による波浪特性の違いを考慮し、日本海に面する留萌・ 浜田、太平洋に面する常陸那珂・細島、瀬戸内海に面する神戸の5つの地点



図3:予測対象とした5つの地点

で波高予測を行った(図 3). 図 4 は常陸那珂における予測値と観測値の 誤差の度数分布を 27 ケース分まとめて示したものである. 裾が負方向に 少し長いが,山の頂点は 0 付近にあり概ね観測値に近い予測値を出せて いる. さらに表 1 には, 5 つの地点において DNN による予測値と WAM を用いた数値計算による予測値の RMSE の 27 ケースの平均を比較したも のを示す.場所によって精度が異なり,留萌・浜田・神戸・常陸那珂にお いては DNN の方が精度が高いが,細島においては数値計算の方が高精度 であった.また図には示していないが,各地点においてケースによる精 度のばらつきが見られた.

4. ベイズ推定による最適推定値の導出

複数の値を使って分散を最小にするように1つの最適値を決める推定 方法である線形最小分散推定(ベイズ推定の一種)を用いた.各推定値 の誤差相関とバイアスがないことを仮定した条件の下で推定式は式(1) のように表せる.

$$x_a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} x_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} x_2 \qquad \dots (1)$$

ここで、 x_1, x_2 は推定値、 σ_1^2, σ_2^2 は真値からの推定値が持つ誤差分散、 x_a は最適推定値である.本研究ではデータ間に誤差相関とバイアスがあることを確認したため、式(2)のように理論的に式変形を行い、これを用いて最適な波高予測値を得た.ここでrは相関係数、 η はバイアスである.

$$x_a = \frac{\sigma_2^2 - r\sigma_1 \sigma_2}{\sigma_1^2 + \sigma_2^2 - 2r\sigma_1 \sigma_2} (x_1 - \eta_1) + \frac{\sigma_1^2 - r\sigma_1 \sigma_2}{\sigma_1^2 + \sigma_2^2 - 2r\sigma_1 \sigma_2} (x_2 - \eta_2) \qquad \cdots (2)$$

また逐次合理性が成り立つことから、2 つの予測データから求めた最 適推定値と別の予測データを使って、新たな最適推定値を求められる. 27 ケースのデータに対して逐次的にこの操作(同化)を行った.図5 は細島における同化に伴う RMSEの推移を示しており、ベイズ推定に よってより高精度な予測値を導出できていることが分かる.それぞれ 青点が DNN による予測値のみを用いた同化、緑点が数値計算による 予測値のみを用いた同化、赤点が DNN と数値計算の両方を用いて同



表 1: DNN と数値計算の RMSE の比較

	DNN	数値計算
留萌	0.585	0.633
浜田	0.522	0.525
常陸那珂	0.566	0.654
神戸	0.121	0.196
細島	0.498	0.465



図 6:同化後の予測時系列

化したものである.2回目の同化以降の変化が小さいのはデータを追加しても新たな情報量が減ってくるためと考えられる.また図6は細島における,DNNと数値計算の両方の予測値を用いた同化によって導かれた最適推定値の 予測時系列の一部である(赤実線,RMSE=0.347).27ケースそれぞれのDNNによる予測値(破線,RMSE=0.436~0.528) と比較すると,同化前に比べて同化後のRMSEが大きく減少しており,観測値により近い推定値が得られた.

5. おわりに

DNN とベイズ推定を併用することで現実的な計算コストで、複数の予測値から、さらに精度の良い1つの予測値 を導く波浪予測システムを構築した.細島や神戸の波高予測では高精度な予測値が得られた.なお波高と同様に、 周期・波向に関しても本システムを適用した結果を発表時に紹介する.

参考文献

1) 齋藤ら(2016): ニューラルネットワークを用いた日本海沿岸域でのうねり性高波浪の予測に関する研究,土木学会論文集 B2 (海岸工学), Vol.72, No.2, L_175-L_180.