

# 機械学習を用いた短中期波浪予測システムの構築

水圏防災工学研究室 久々江 耀平  
指導教員 細山田 得三

## 1. 序論

我が国における沿岸域は、海底ケーブルや洋上風力発電、貿易、レジャーなど経済上の重要な役割を担う。近年の水災害の激甚化も相まって沿岸域の防災・減災がより一層求められており、その第一段階として波浪予測技術の向上が喫緊の課題とされている。現在では、気象庁の全球波浪数値予報モデル GPV (GWM)による 1 週間波浪予測等を元に沿岸域の工事、船舶運航、レジャー等の可否判断を行なっている<sup>1)</sup>。近年、新たな予測技術として機械学習が注目され、物理モデルに近い予測精度を持つ事例もある。

Tom らの事例<sup>2)</sup>では、GWM 等の全球モデルによる予測の空間解像度を改善するために機械学習モデル Gradient Boosting Decision Tree (GBDT) の実装の一つである XGBoost を用いた結果、GWM 以上の精度を持つ波浪予測システムとなった。具体的には、GWM の予測値の生データを特徴量、那覇港と中城港の NOWPHAS の有義波・波高観測値を教師データとしてモデルを構築しており、Tom らは機械学習を用いてある点のデータを任意の点におけるデータへの変換を行なったと述べている。

こうした機械学習を用いた波浪予測事例の共通した課題として機械学習手法の独自化が挙げられる。他分野の機械学習においては、モデルに有効・有意とされる説明変数（特徴量）を生データから作成・選択する特徴量エンジニアリングが不可欠である。また、XGBoost をはじめとして機械学習は予測の根拠を明示することのための手法があるが、それを用いてモデルを説明した例はない。従って、機械学習を用いた波浪予測は多分に改善の余地がある。

以上から本研究では、機械学習を用いた 1 週間先まで波浪予測システムを構築するとともに、他分野で一般的とされる手法を実践することで波浪予測における機械学習の妥当な評価・検討をする。このシステム構築に当たり、留意すべき事項として、低波浪時かつ 1 週間先までの波高を回帰分析・多変量時系列解析的に予測すること、入出力をテーブルデータとして扱うこと、波浪予測における特徴量エンジニアリングの最適手法を模索することを示す。特に本研究のシステムは海上工事の作業可否判断の元となる 1m から 1.2m の荒天日波高閾値を対象とする。

## 2. 使用モデル

本研究で使用した機械学習モデルは Gradient Boosting Decision Tree (GBDT) と TabNet である。

### XGBoost

Decision Tree はある条件（例：波高の過去 10 日間の標準偏差が 1 位下）が設定された複数のノードから成る木構造の計算グラフで表される。最初に特徴量を受け取るノードを根、値の伝播を枝、その先のノードを葉という。各葉では条件に応じた真偽が返され次の葉と予測値を形成する。この葉の条件を学習により適切なものに更新し、予測を行う。

GBDT は Decision Tree をベースに勾配降下法、アンサンブル学習を合わせた機械学習モデルである。テーブルデータに有効とされる。本研究では、その実装ライブラリの一つである XGBoost<sup>3)</sup>を用いる。XGBoost をはじめとした GBDT の実装ライブラリには、モデルが各特徴量を重視する程度を示す Feature Importance 算出機能が備わっていることが挙げられる。Feature Importance は各特徴量が Decision Tree の分岐の条件として採択された回数を本研究では、大量に作成した特徴量を GBDT に入力し、Feature Importance を算出することで特徴量の要不要を判断する。

### TabNet

TabNet<sup>4)</sup>は Neural Network を用いてを模した深層学習モデルである。TabNet は深層学習モデルでありながら GBDT と同様にテーブルデータに有効であるとされ、Feature Importance の算出機能を持つ。一方で TabNet は他深層学習モデルと同様に特徴量作成の影響度が低いことが報告されている。

## 3. 実験

本研究のシステムでは 2 時間ごとあるいは 6 時間ごとに更新される波浪観測値及び気象解析値から選択済み特徴量を作成、各学習済みモデルへ入力、予測値の算出を行う。そのための特徴量作成、学習、特徴量選択、再学習について述べる。

## 特徴量作成

NOWPHAS<sup>9)</sup>の有義波・波高観測値と JRA-55<sup>6)</sup>の  $u \cdot v$  方向風速解析値から特徴量を作成し、有義波・波高は教師データとしても使用する。NOWPHAS の波浪観測点・期間は、那覇港 (2010 年から 2019 年) と新潟港 (2014 から 2018 年) を含む日本海側の観測点とする。那覇港は Tom らの事例との精度比較のために用いている。新潟港含む日本海側の複数観測点は複数の波高データの影響を知るために用いている。風速データはより古いデータからも有効な傾向を見出し、ノイズとしないという点から JPA-55 を選出した。

特徴量作成の内、対数変換、正規化、年月日時の変数変換はモデル内で扱う数値のスケールを揃える意図で行なっている。また、基本統計量 (標準偏差, 歪度, 尖度), 年月日時の変数変換, ラグ特徴量はモデルにデータの時系列的依存を知らせる意図で作成している。

これらの特徴量を学習データ (~2016/12/31), 検証データ (2017/1/1 ~ 2017/12/31), 評価データ (2018/1/1 ~) に分けモデルの学習・評価を行う。

## 特徴量選択

学習後, XGBoost の予測時の Feature Importance 値が 10 以上, 7 以上, 5 以上, 3 以上, 1 以上の組み合わせを選択し。再学習・再評価を行い, さらにモデルを最適化する。その結果, 精度が最も向上した特徴量の組み合わせを最終的なシステムへの入力とする。

## 4. 結果と考察

### 特徴量について

実験の結果, TabNet, XGBoost は赤道付近の風速データの統計量を重視していることが明らかになった。これは物理モデルと同様に物理法則のような因果関係を表している。従って, 機械学習モデルは学習により物理法則を近似しているといえる。

また, このことから機械学習に影響を与える風速データなどを 1 つにまとめた高度な交互作用特徴量が GWM による予報値であるといえる。それ故に Tom らによる事例<sup>2)</sup>は特徴量エンジニアリングを実行せずとも既往研究以上の精度を持つに至ったと推察される。

本研究においては GWM を用いた場合, 高波浪時へのフィッティングが改善されると考えられるが, 同時に低波浪時におけるノイズとなり得る。故に高波浪を対象と

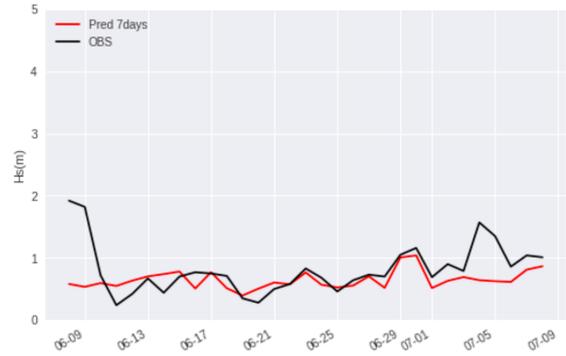


図-1 2019/6/8-7/8 那覇港 1週間先有義波高(m)

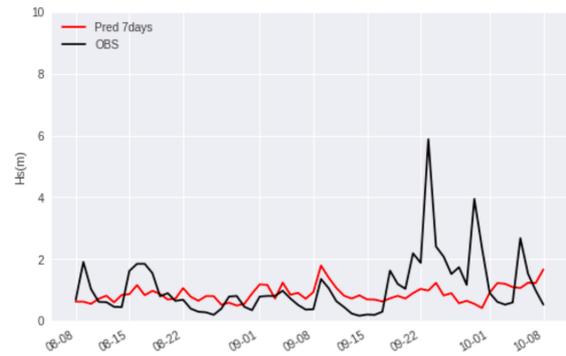


図-2 2018/8/8~10/8 那覇港 1週間先有義波高(m)

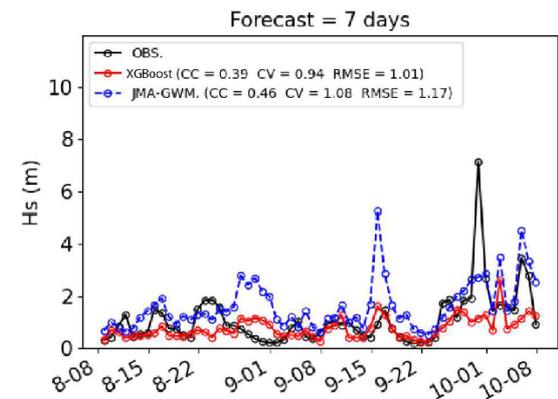


図-3 2018/8/8~10/8 那覇港 1週間先有義波高(m)

Tom らによる事例<sup>2)</sup>より

した別モデルを作成する必要がある。

### 予測精度について

図-1 に示すように那覇港の 1m 以下の低波浪時は, 比較的フィッティングが良いが, 高波浪時には, 変化の傾向を補足しきれていないことがわかる。同様の傾向が新潟港, 他期間でも見られた。

また, 図-2 に本研究での 1 週間予測の結果を示し, 同期間の Tom らによる予測結果を図-3 に示す。これらと比較すると全体的な変化の傾向はおおよそ捉えていると

いえる。しかし、図-3 では観測値ではなく、特徴量とした GWM の傾向に追従した予測をするのみである。

本研究の 1 週間予測における RMSE (Root Mean Squared Error) は、Tom らの 1.01(m) から 3 割改善し、0.70(m) であった。これは先述の特徴量の差異に起因すると推察される。

## 5. 結論

本研究では、機械学習により海岸工事の可否判断に用いる低波浪時の 1 週間先までの波浪予測を行った。特徴量エンジニアリング、モデル解釈、予測の結果、低波浪時の 1 週間予測において最も良い性能を持つといえる。また、1 週間先の荒天日波高閾値を予測可能なシステムである。

謝辞：本研究を行うに当たり、NOWPHAS データを使用致しました。関係者の皆様には波浪観測結果の利用に關しまして感謝の意を表します。

## 参考文献

- 1) 片山 裕之, 前田 勇司: 港湾工事の供用係数に関する一考察: 土木学会論文集 B3 (海洋開発) vol.75, pp1\_923-I\_928: 2019
- 2) Tracey H. A. Tom, 間瀬 肇, 池本 藍, 川中 龍児, 武田 将英, 原 知聡, 金 洙列: GWM と XGBoost を用いた 1 週間波浪予測: 土木学会論文集 B3 (海洋開発) vol.77, pp1\_7-I\_12: 2021
- 3) Tianqi Chen, Carlos Guestrin: XGBoost: A Scalable Tree Boosting System: <https://arxiv.org/abs/1603.02754>: 最終アクセス 2022/2/15
- 4) Sercan O. Arik, Tomas Pfister: TabNet: Attentive Interpretable Tabular Learning: <https://arxiv.org/abs/1908.07442>: 最終アクセス 2022/2/15
- 5) 国土交通省港湾局 全国港湾海洋波浪情報網: リアルタイムナウファース: <https://www.mlit.go.jp/kowan/nowphas/index.html>: 最終アクセス 2022/2/5
- 6) JRA-55 project: [https://jra.kishou.go.jp/JRA-55/index\\_ja.html](https://jra.kishou.go.jp/JRA-55/index_ja.html): 最終アクセス 2022/2/5