

# GWM と XGBoost を用いた1週間波浪予測

- ▶ Tracey H. A. Tom (株)ハイドロ総合技術研究所
- ▶ 間瀬 肇 京都大学名誉教授
- ▶ 池本 藍 HOKUULA Corp.
- ▶ 川中 龍児 (株)ハイドロ総合技術研究所
- ▶ 武田 将英 東亜建設工業(株)
- ▶ 原 知聡 東亜建設工業(株)
- ▶ 金 洙列 熊本大学准教授



# まえがき<sub>1</sub>

- ▶ 特別セッション: S1.「働き方改革を踏まえた港湾・海岸事業の進め方」. 3つのテーマのうちの1つが「作業可否の事前予測」.  
→ 特別セッションに投稿
- ▶ 大規模海洋工事の例として洋上風力発電装置の設置がある. 基礎構造物の施工, タワーの据付け, 風車本体の施工, 海底ケーブルの敷設といった工種の中には, 数日間連続して作業を実施しなければならず, おおむね1週間にわたる波浪の予測が必要となる.



## まえがき<sub>2</sub>

- ▶ 人工知能AIは、ビッグデータを処理して、汎用的には人間の感性や判断能力を持たせる技術、特化的にはある特定の作業において人と同等かそれ以上の処理能力を持たせる技術。
- ▶ 著者らは、単独あるいは複数の空間解像度の粗い全球波浪予報値を GMDH や ANN を用いて、日本沿岸の波浪予測値として利用する波浪予測法を提案してきた。  
→ 実運用は今年中
- ▶ 特化型の機械学習は、反復的に学習を行いパターンや特徴を見つけ出して、予測を行うものである。
- ▶ 本研究は、アンサンブル学習の1つである XGBoost(勾配ブースティング法)を利用して、全球波浪モデル GWM 予報値を、日本沿岸の波浪値に変換する波浪予測モデルを構築し、その予測精度を検討する。



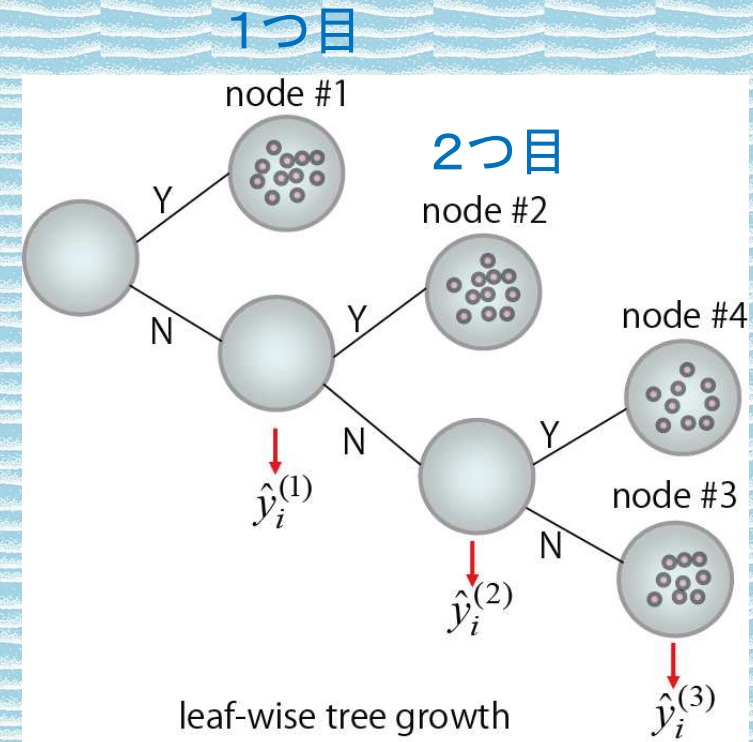
# 決定木とアンサンブル学習

- ▶ 決定木は、分類木と回帰木を組み合わせたもので、ツリーによってデータを分析する手法
  - ▶ 各要素に対して、区分結果を表したツリーが「分類木」
  - ▶ 連続して変わりうる値を分析する場合には「回帰木」
- ▶ アンサンブル学習は、複数のモデルを同時に学習し、それらを相互に用いることで1つの学習モデルとしてモデルの精度を向上させる手法
  - ▶ バギング: ブートストラップサンプリングによって得た学習データを用い、複数の決定木を作って多数決
  - ▶ ランダムフォレスト: 複数データに対して、同じ構造の決定木にならないように特徴量をランダムに与えて複数の決定木を作って多数決
  - ▶ XGBoost: バギング, ランダムフォレストは並列に, XGBoostは直列的に決定木を作成. 最初の学習器で上手く推定できなかった部分(誤差)を推定するために次の学習器で学習



# XGBoostの定式化<sub>1</sub>

- ▶ 1つ目の決定木を1つ構築し、この決定木の入力に対する予測値の誤差を求め、この誤差を目的変数として2つ目の決定木を構築する。
- ▶ 2つ目の決定木の予測値は、誤差の予測値を1つ目の予測値に加えるという手順を繰り返して、目的値と予測値の誤差が最小になるようにする。
- ▶ 1つ前の決定木が残した誤差を、次の決定木で補正しながら学習精度を向上させていくBoosting.





## XGBoostの定式化<sub>2</sub>

- ▶  $k$ 個目の決定木による予測関数を  $f_k$  とする.
- ▶ 学習時に用いる損失関数は、予測値と観測値の2乗誤差、決定木を構築したときの最終ノード数、予測値の集合(ノルム)の和からなる。また、損失関数は分析者が決めるパラメータを含み、最終ノード数や更新のスピードを変えることができる。
- ▶  $t$  番目の決定木ノードを右  $R$  と左  $L$  に分岐するとき、分岐前と分岐後の損失関数の値が小さくなるように分岐する。



# XGBoost の応用

時間があれば最  
後に紹介

## Gradient Boosting の適用によるダム流入量予測の精度向上

天方 匡純<sup>1</sup> 藤井 純一郎<sup>2</sup> 梁田 信河<sup>2</sup>

## ダム流入量予測

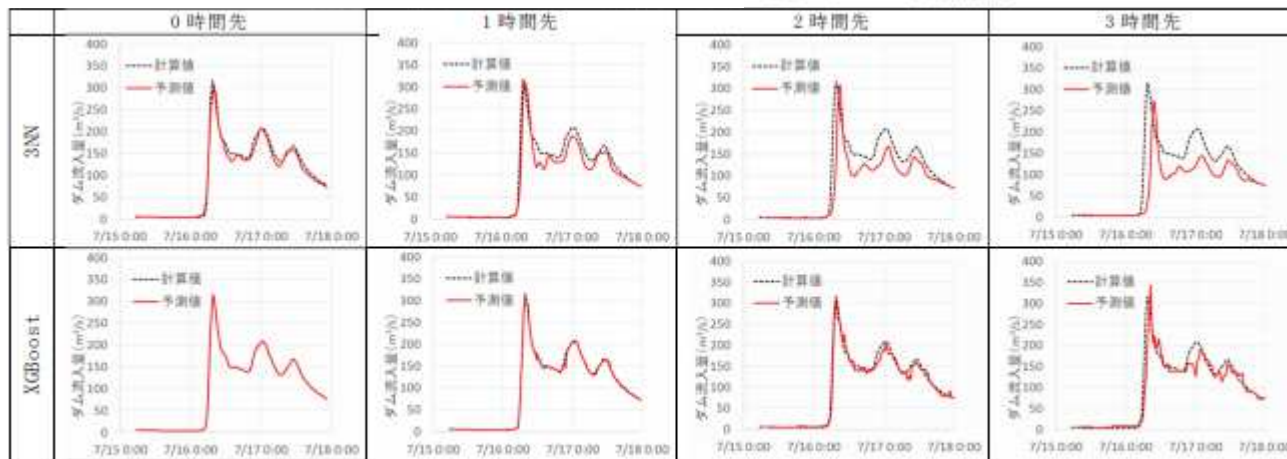
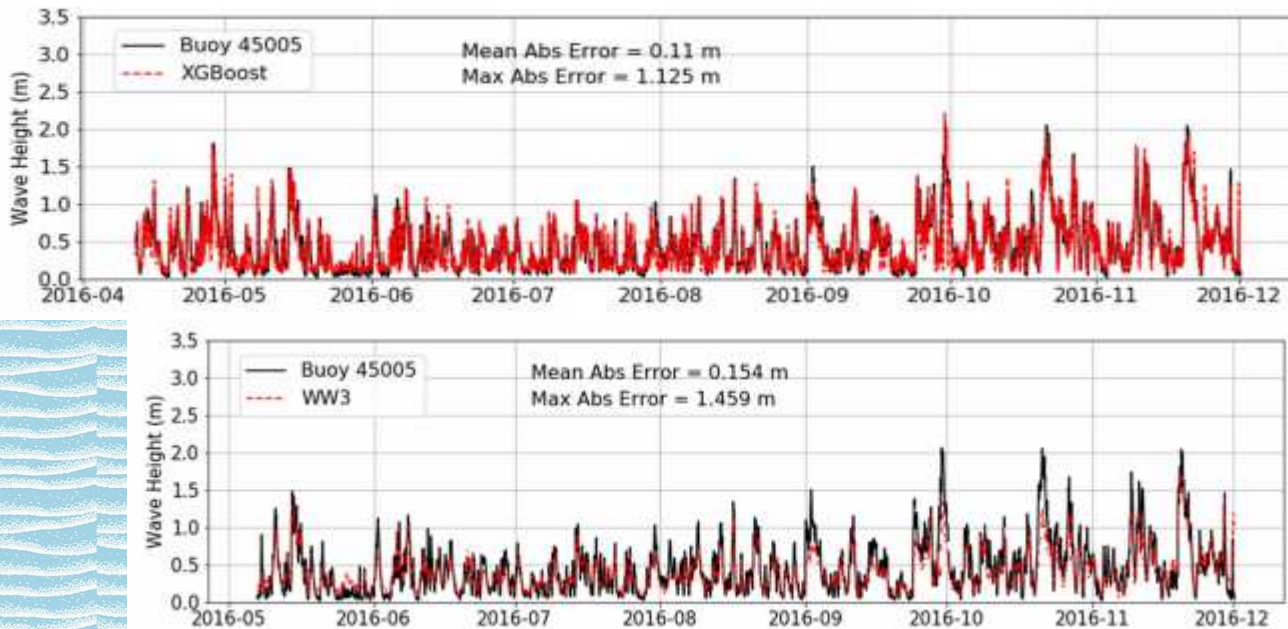


図-10 平成 27 年 7 月 16 日出水に対する各モデルの予測状況

## Predicting Lake Erie Wave Heights using XGBoost

Haoguo Hu<sup>1</sup> and Philip Chu<sup>2</sup>  
1. University of Michigan, Ann Arbor 2. NOAA/GLERL

## 波浪予測





# 全球波浪GWM予報値とXGBoostを用いる波浪予測<sub>1</sub>

- ▶ GWM は外力として用いる気象の予測に誤差があり、また空間解像度は粗く(0.5度×0.5度)沿岸の波浪変形を考慮しないため、観測値と合わない。
- ▶ GWM 予報値 {波高(m), 周期(s), 波向(度)} は、初期値を 00, 06, 18 UTC として、予報時間は 6時間間隔で 132 時間までである。初期時刻が 12時UTC の場合には 6時間間隔で 138 ~ 264 時間の予報値が提供されている。
- ▶ 本研究では、毎日 12:00 時 (UTC) に発表される 1, 3, 5, 7 日先のGWM予報値から XGBoost を用いて予測値を求め、それぞれの時刻での波浪観測値と比較する。



## 全球波浪GWM予報値とXGBoostを用いる波浪予測<sub>2</sub>

- ▶ 入力として与えるデータは, GWM予報値である波高  $H_s$  と周期  $T_s$ , および波の方向成分  $U, V$  の 4 つである.
- ▶ 教師データは, 予報対象時刻の観測波高である. 入力データに対して XGBoost により観測波高  $H_s$  に近い値が出力できるように決定木を決定していく.
- ▶ XGBoost のハイパーパラメータの決定には, Optunaのベイジアン最適化法を用いる.
- ▶ 対象とした地点は, 那覇港および中城港である. 期間は, 2010年1月1日から2018年12月31日の9年間の, 毎日の12時UTCにおける値と 1, 3, 5, 7日先のデータである.
- ▶ 2010年から 2017年のデータのうち 80%を訓練に, バリデーションデータは残りの 20%, テストデータは 2018 年のデータとした.



# 観測結果と予測結果の比較検討<sub>1</sub>

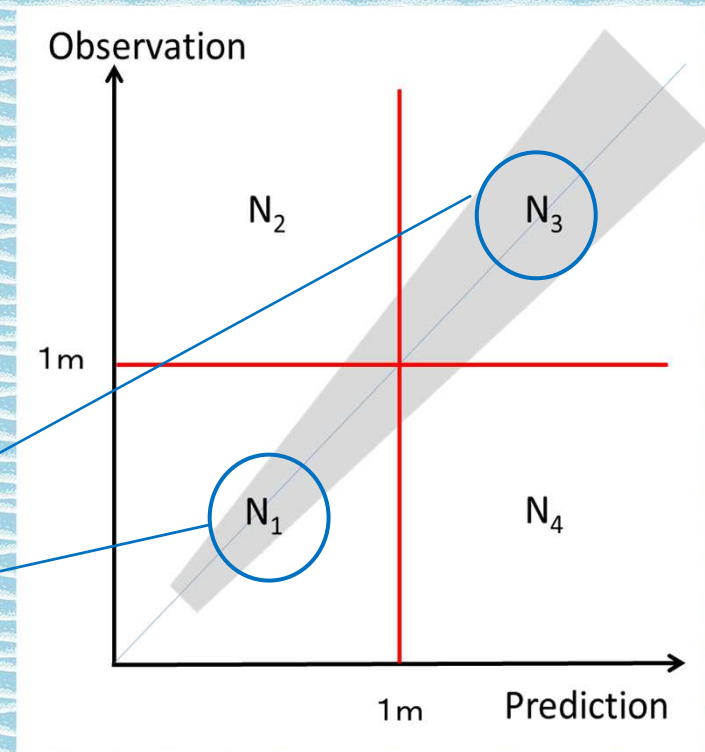
## ▶ 用いた指標

$$CC = \frac{\sum_{i=1}^n (o_i - \bar{o})(p_i - \bar{p})}{\sqrt{\sum_{i=1}^n (o_i - \bar{o})^2} \sqrt{\sum_{i=1}^n (p_i - \bar{p})^2}}$$

$$CV = \frac{\text{RMSE of Obs. and Pred. Values}}{\text{Mean Value of Observations}}$$

$$HIT_1 = \frac{N_1 + N_3}{N_1 + N_2 + N_3 + N_4}$$

$$HIT_2 = \frac{N_1}{N_1 + N_2}$$



波高を1mより大きい、あるいは、  
1mより小さいと予測した時の的中率



# 観測結果と予測結果の比較検討<sub>1</sub>

## ▶ 用いた指標

$$CC = \frac{\sum_{i=1}^n (o_i - \bar{o})(p_i - \bar{p})}{\sqrt{\sum_{i=1}^n (o_i - \bar{o})^2} \sqrt{\sum_{i=1}^n (p_i - \bar{p})^2}}$$

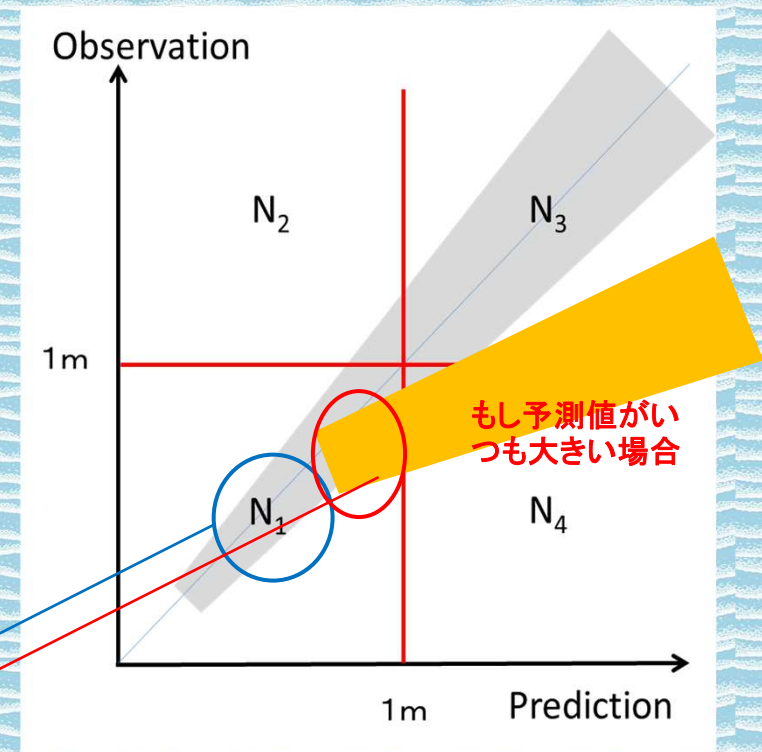
$$CV = \frac{\text{RMSE of Obs. and Pred. Values}}{\text{Mean Value of Observations}}$$

$$HIT_1 = \frac{N_1 + N_3}{N_1 + N_2 + N_3 + N_4}$$

1mより小さいと予測した時の的中率

$$HIT_2 = \frac{N_1}{N_1 + N_2}$$

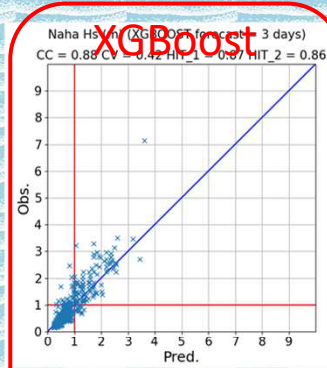
予測値が観測値より、いつも大きいと的中率 HIT<sub>2</sub> は1になる (N<sub>2</sub>=0)



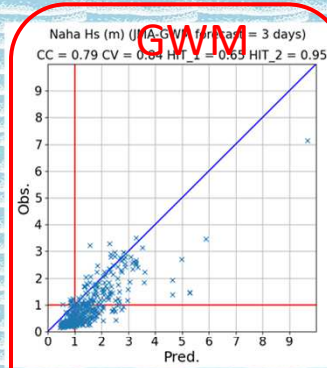


# 観測結果と 予測結果の 比較検討<sub>2</sub>

## 那覇港

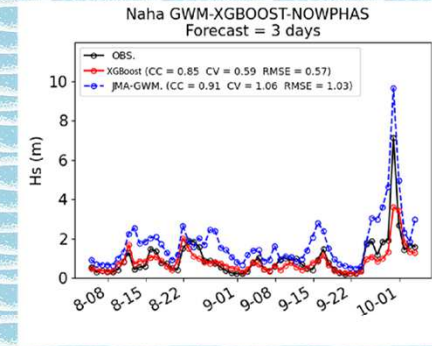


XGBoost モデル

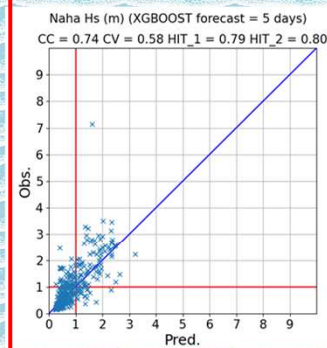


GWM

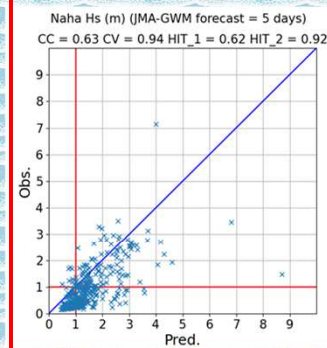
(a) 3日先予測



2ヶ月間の予測値と観測値の時系列比較

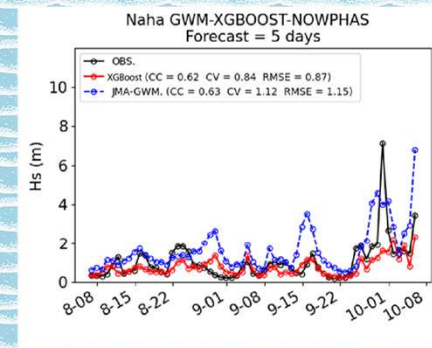


XGBoost モデル

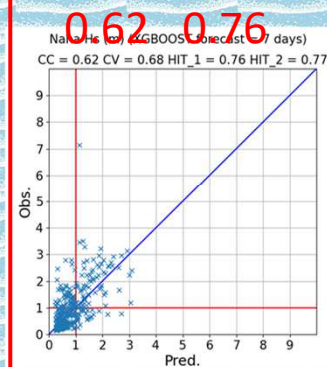


GWM

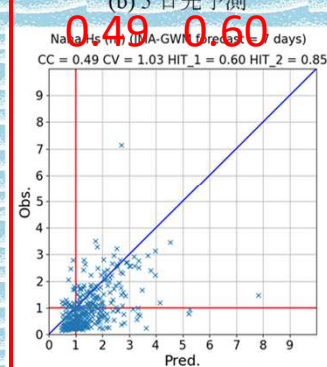
(b) 5日先予測



2ヶ月間の予測値と観測値の時系列比較

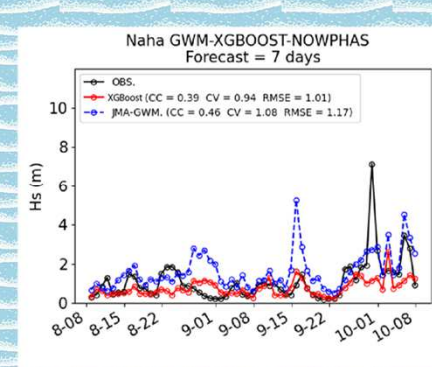


XGBoost モデル



GWM

(c) 7日先予測



2ヶ月間の予測値と観測値の時系列比較

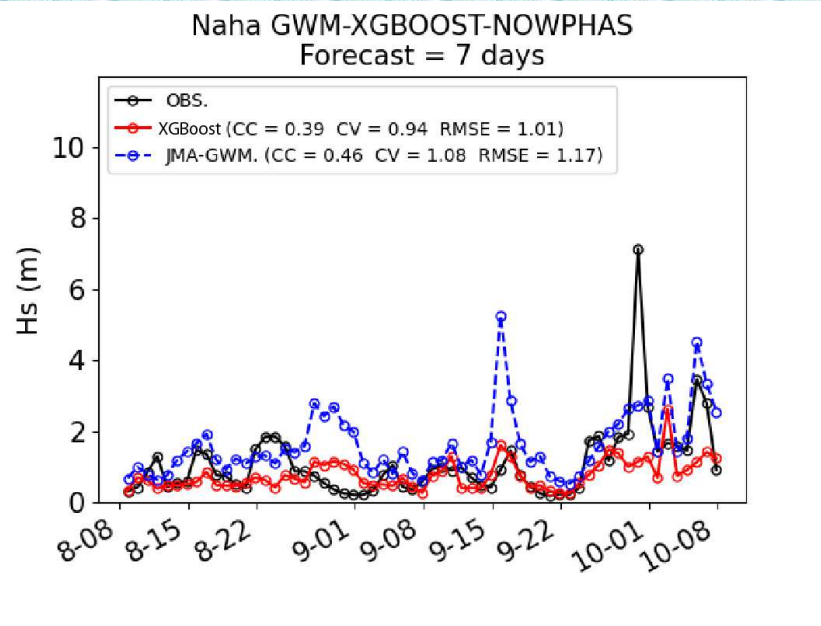
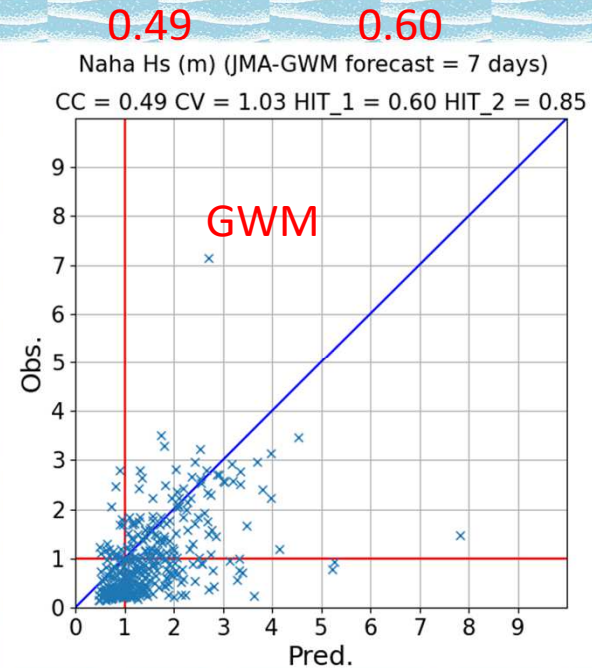
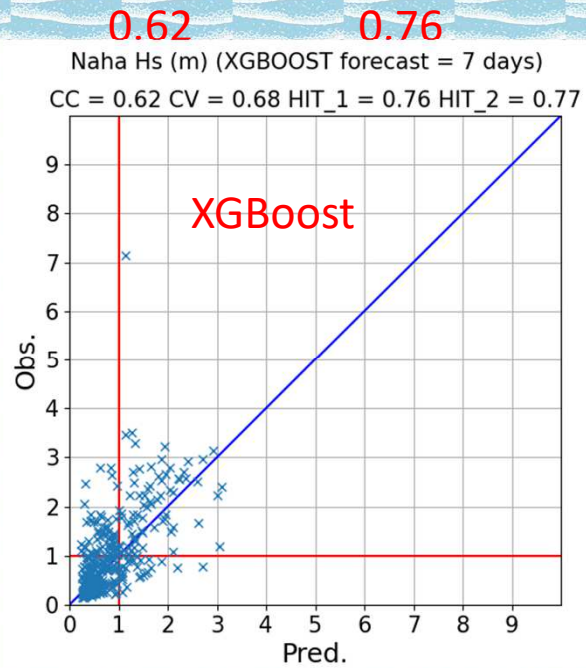
次ページに大きくする

図-3 那覇港の波高に関する XGBoost 予測値, GWM 予報値および観測値の比較



# 7日先

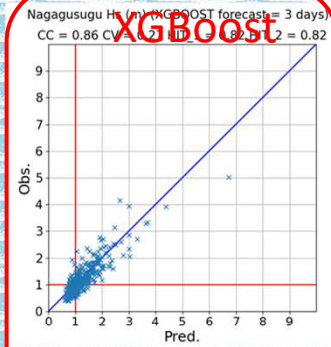
# 那覇港



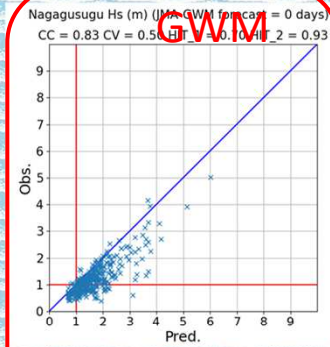


# 観測結果と 予測結果の 比較検討<sup>3</sup>

## 中城港

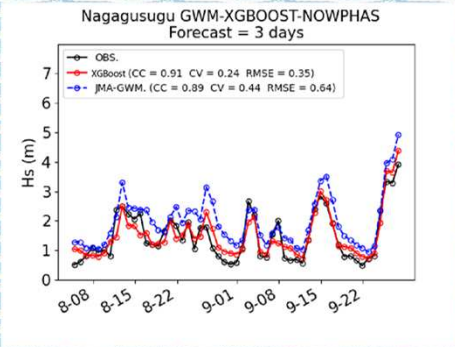


XGBoost モデル

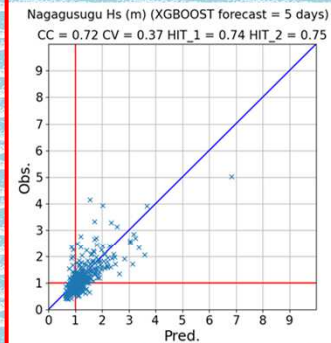


GWM

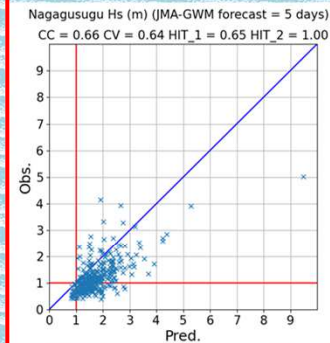
(a) 3日先予測



2ヶ月間の予測値と観測値の時系列比較

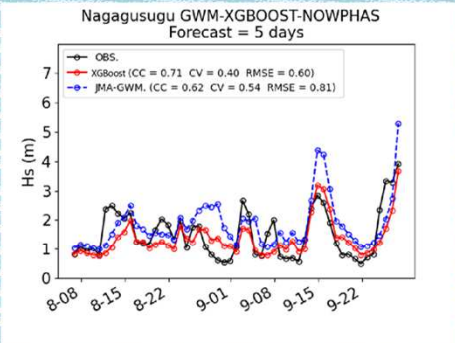


XGBoost モデル

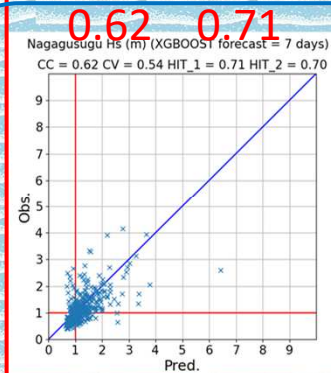


GWM

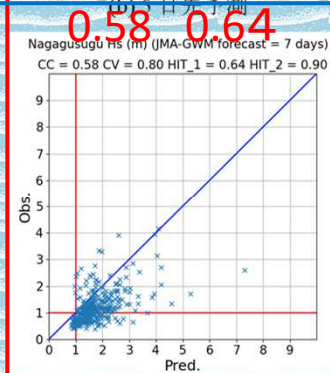
(b) 5日先予測



2ヶ月間の予測値と観測値の時系列比較

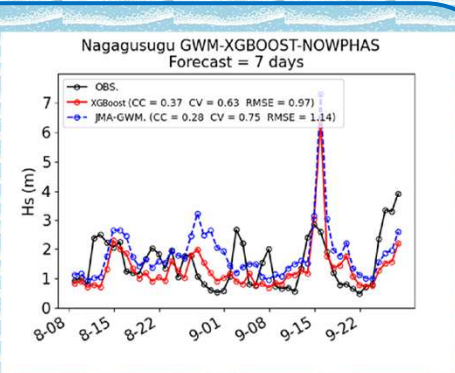


XGBoost モデル



GWM

(c) 7日先予測



2ヶ月間の予測値と観測値の時系列比較

次ページに大きくする

図-4 中城港の波高に関する XGBoost 予測値, GWM 予報値および観測値の比較



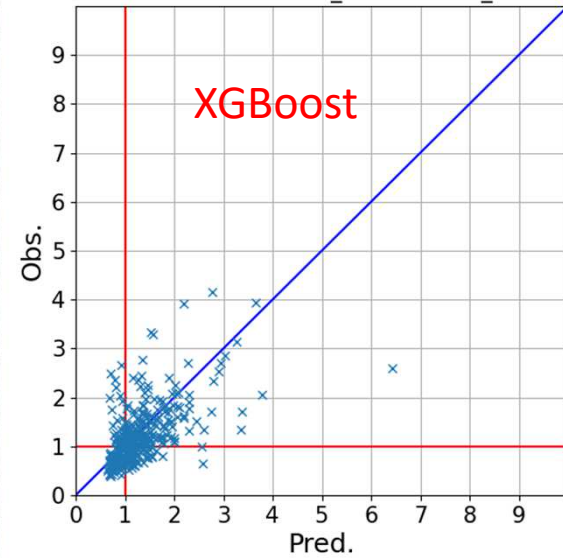
# 7日先

# 中城港

0.62

0.71

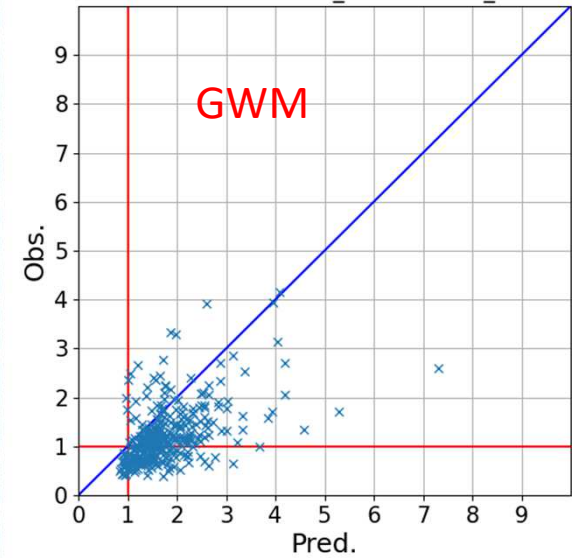
Nagagusugu Hs (m) (XGBOOST forecast = 7 days)  
CC = 0.62 CV = 0.54 HIT\_1 = 0.71 HIT\_2 = 0.70



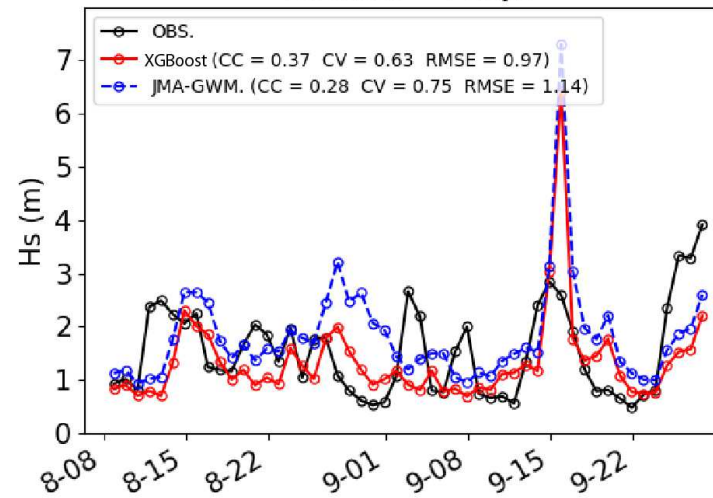
0.58

0.64

Nagagusugu Hs (m) (JMA-GWM forecast = 7 days)  
CC = 0.58 CV = 0.80 HIT\_1 = 0.64 HIT\_2 = 0.90



Nagagusugu GWM-XGBOOST-NOWPHAS  
Forecast = 7 days





# 観測結果と予測結果の比較のまとめ

- ▶ 那覇港および中城港において、7日先予測において波高が1 m 以下か、以上かを予測したときの的中率は GWM 予報値では 0.7 以下であるが、XGBoost 予測値は 0.7 以上となり、予測精度が良くなることがわかった。
- ▶ また、XGBoost の3日先予測値の的中率は 0.82 であり、原らが解析した沿岸波浪予報値CWMの那覇港における的中率の 0.88 と比べて少し悪くなるが、GWMの 0.70 より良い。



# まとめ

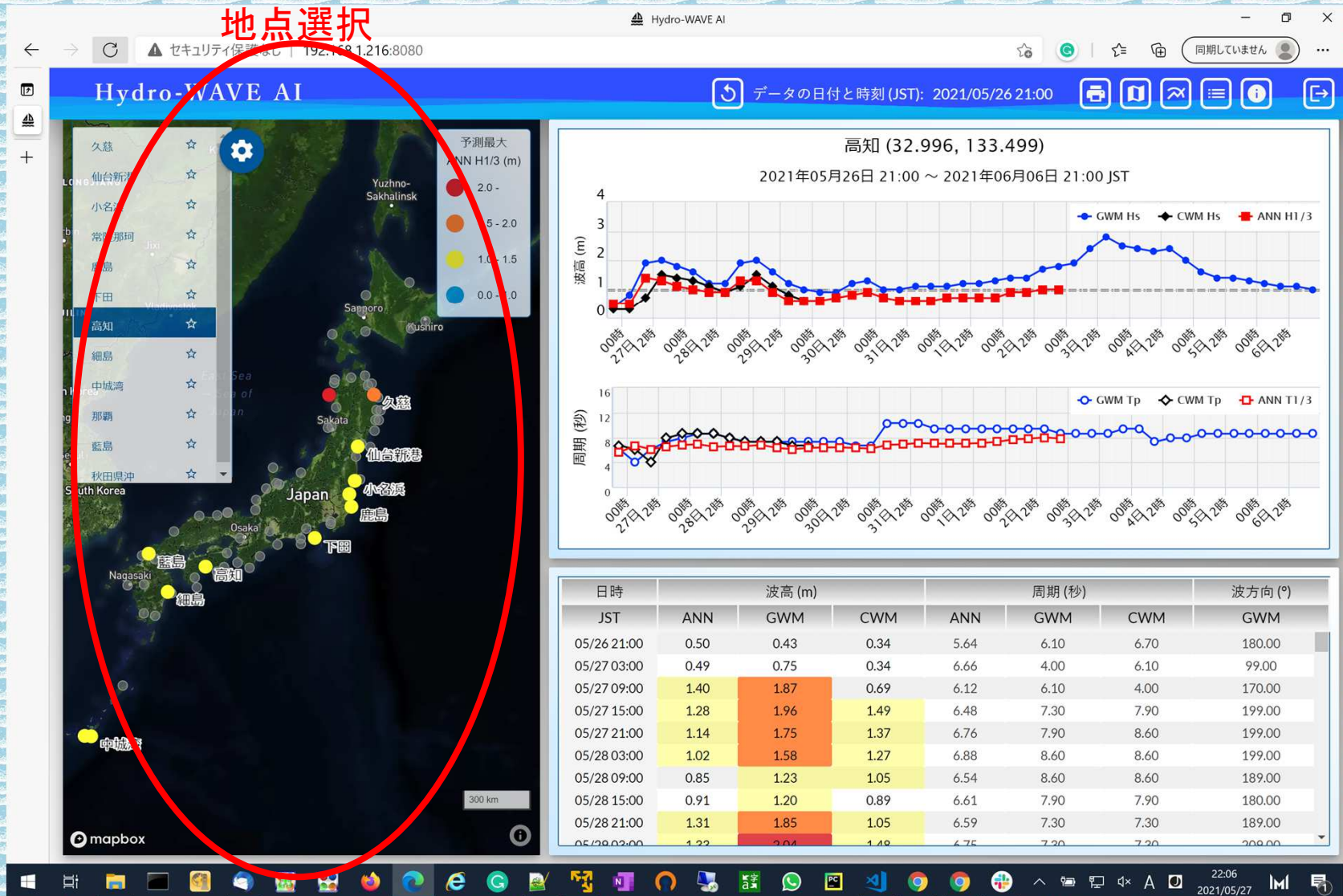
- ▶ 全球波浪予報は日本沿岸の波浪予測値として利用するには空間解像度が低いので、機械学習を利用して1週間先までの波浪予測を行う GMDH 手法や ANN 手法を利用する波浪予測モデルが提案されている。
- ▶ 本研究では、アンサンブル学習手法の1つである XGBoost を利用して、全球波浪予報値 GWM を日本沿岸の波浪予測値に変換する GWM to XGBoost 波浪予測モデルをつくり、その予測精度を検討した。
- ▶ 那覇港および中城港において、7日先予測において波高が1m以下か、以上かを予測したときの的中率は、予報値では0.7以下であるが、XGBoost 予測値は観測値と一致となり予測精度が良くなること、7日先予測値と観測値との相関係数は0.6以上であった。

実際の運用は？



# 1週間波浪予測システム

地点選択





# 1週間波浪予測システム

Hydro-WAVE AI

データの日付と時刻 (JST): 2021/05/26 21:00

図とデータを  
1ページ印刷

関連情報

高知 (32.996, 133.499)  
2021年05月26日 21:00 ~ 2021年06月06日 21:00 JST

日時 JST	波高 (m)			周期 (秒)			波方向 (°)
	ANN	GWM	CWM	ANN	GWM	CWM	GWM
05/26 21:00	0.50	0.43	0.34	5.64	6.10	6.70	180.00
05/27 03:00	0.49	0.75	0.34	6.66	4.00	6.10	99.00
05/27 09:00	1.40	1.87	0.69	6.12	6.10	4.00	170.00
05/27 15:00	1.28	1.96	1.49	6.48	7.30	7.90	199.00
05/27 21:00	1.14	1.75	1.37	6.76	7.90	8.60	199.00
05/28 03:00	1.02	1.58	1.27	6.88	8.60	8.60	199.00
05/28 09:00	0.85	1.23	1.05	6.54	8.60	8.60	189.00
05/28 15:00	0.91	1.20	0.89	6.61	7.90	7.90	180.00
05/28 21:00	1.31	1.85	1.05	6.59	7.30	7.30	189.00
05/29 03:00	1.22	2.04	1.48	6.75	7.20	7.20	200.00

Windows taskbar: 22:06 2021/05/27