

はじめに

GMDH

データ

テスト

結果

まとめ

論文紹介

Development of GMDH-Based Storm Surge Forecast Models for Sakaiminato, Tottori, Japan (GMDH に基づく 鳥取県境港市の 高潮予測モデルの開発)

中市新太

富山県立大学

u020025@st.pu-toyama.ac.jp

背景

高潮による被害により沿岸部では迅速な意思決定を行う必要がある．早期警報システムが必要になる

先行研究

1. プロセスベースの数値予測モデル
浅水方程式からなる二次元モデル．計算負荷が軽い．
2. 機械学習ベースの予測モデル
人口ニューラルネットワーク (ANN), ディープラーニング
学習データに教師される必要がある．

GMDH 予測モデル

x_i, x_j は2つの入力, y_n は出力, n は整数であるとして (1) を計算する.

$$y_n = a_{k_0} + a_{k_1} x_{i_n} + a_{k_2} x_{j_n} + a_{k_3} x_{i_n}^2 + a_{k_4} x_{j_n}^2 + a_{k_5} x_{i_n} x_{j_n} \quad (1)$$

y_{on} を観測値として,

$$g = \sum_{n=1}^N (y_n - y_{on})^2 \quad (2)$$

$$\frac{\partial g}{\partial a_k} = 0 \quad (3)$$

によって a_k を求める.

GMDH 予測モデル

(3) を展開することで (4) から (9) のガウス正規方程式が得られる。

$$\sum_{n=1}^N y_{kn} = \sum_{n=1}^N \{a_{k0} + a_{k1}x_{in} + a_{k2}x_{jn} + a_{k3}x_{in}^2 + a_{k4}x_{jn}^2 + a_{k5}x_{in}x_{jn}\} \quad (4)$$

$$\sum_{n=1}^N y_{kn}x_{in} = \sum_{n=1}^N \{a_{k0}x_{in} + a_{k1}x_{in}^2 + a_{k2}x_{jn}x_{in} + a_{k3}x_{in}^2x_{jn} + a_{k4}x_{jn}^3 + a_{k5}x_{in}x_{jn}^2\} \quad (5)$$

$$\sum_{n=1}^N y_{kn}x_{jn} = \sum_{n=1}^N \{a_{k0}x_{jn} + a_{k1}x_{jn}x_{in} + a_{k2}x_{jn}^2 + a_{k3}x_{in}^2x_{jn} + a_{k4}x_{in}^2x_{jn} + a_{k5}x_{in}^3\} \quad (6)$$

$$\sum_{n=1}^N y_{kn}x_{in}x_{jn} = \sum_{n=1}^N \{a_{k0}x_{in}x_{jn} + a_{k1}x_{in}^2x_{jn} + a_{k2}x_{in}x_{jn}^2 + a_{k3}x_{in}^2x_{jn}^2 + a_{k4}x_{in}^3x_{jn} + a_{k5}x_{in}x_{jn}^3\} \quad (7)$$

$$\sum_{n=1}^N y_{kn}x_{in}^2 = \sum_{n=1}^N \{a_{k0}x_{in}^2 + a_{k1}x_{in}^3 + a_{k2}x_{in}^2x_{jn} + a_{k3}x_{in}^3x_{jn} + a_{k4}x_{in}^4 + a_{k5}x_{in}^2x_{jn}^3\} \quad (8)$$

$$\sum_{n=1}^N y_{kn}x_{jn}^2 = \sum_{n=1}^N \{a_{k0}x_{jn}^2 + a_{k1}x_{in}x_{jn}^2 + a_{k2}x_{in}^3 + a_{k3}x_{in}x_{jn}^3 + a_{k4}x_{in}^2x_{jn}^2 + a_{k5}x_{in}^4\} \quad (9)$$

図 1: ガウス正規方程式

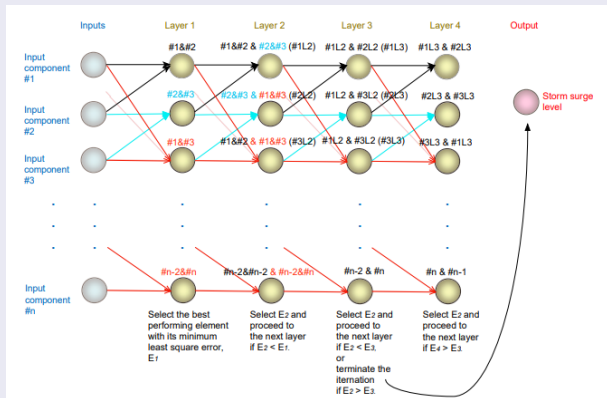


図 2: アルゴリズム

はじめに

GMDH

データ

テスト

結果

まとめ

ステップ

1. 要素のペアを生成
2. (4) から (9) を用いてペアごとの a_k を決定する
3. 各要素について、係数 a_k を用いて y_n を導く
- 4 各要素での予測値と観測値の差である g を計算
5. 要素ごとに g を小さい順に並べ替える
- 6.3 の各要素について 1 から 5 を繰り返す (g が出る)
- 7.5, 6 を比較寿司, 最高精度が前の層から変化しなくなるまで繰り返す

調査地域とデータ

鳥取県の境港は日本海、東海の 2 津の大陸棚に面しており、6 月から 12 月にかけて、台風による高潮の危険にさらされている。1975 年以来、年間 2.6 個の台風が通過または上陸している。

気象データとして、浜田、松江、米子、海士、西郷の 5 つの気象観測所において、「風速」、「風向」、「海面気圧」、「背景気圧 (1013hPa) からの降下率」を収集した。

流体力学データとして、境港の 1 つの流体力学観測点で「海面水位」と「サージ水位」を収集した。

台風情報は、「台風の特徴を表す台風中心付近の位置」、「中心気圧」、「最大風速のデータ」を選定した。

これらのデータは、2003 年の前見台風、2004 年の松田台風、2004 年の女木台風の 3 つの台風で観測された。

データの数値

はじめに

GMDH

データ

テスト

結果

まとめ

$$\eta_i = \tilde{\eta}_i v_{SSL} = \tilde{v}_{SSL} / 2 \text{ m} \quad (4)$$

$$v_{SLP} = \tilde{v}_{SLP} / 1013 \text{ hPa} \quad (5)$$

$$v_{DSL P} = \tilde{v}_{DSL P} / 100 \text{ hPa} \quad (6)$$

$$v_{WS} = \tilde{v}_{WS} / 100 \text{ m/s} \quad (7)$$

$$v_{WD} = \tilde{v}_{WD} / 360 \text{ degree} \quad (8)$$

$$v_{LG} = \tilde{v}_{LG} / 150^\circ \text{ E} \quad (9)$$

$$v_{LT} = \tilde{v}_{LT} / 50^\circ \text{ N} \quad (10)$$

$$v_{CAP} = \tilde{v}_{CAP} / 1013 \text{ hPa} \quad (11)$$

$$v_{HWS} = \tilde{v}_{HWS} / 100 \text{ m/s} \quad (12)$$

(4) サージ水位, (5) 海面水位, (6) 海面気圧, (7) バックグラウンド海面気圧からの海面気圧の低下, (8) 風速, (9) 風向, (10) 台風の経度, (11) 台風の緯度, (12) 台風の中心気圧, (13) 台風中心付近の最大風速

トレーニングテスト

1つの再解析モデルとそれぞれ5時間、12時間、24時間のリードタイムを持つ3つの予測モデルが以下の方程式を用いて開発された。

$$\begin{aligned}\eta_{t=i} &= f(SSL_{t=i}, SLP_{t=i}^{st=1,\dots,5}, DSLP_{t=i}^{st=1,\dots,5}, WS_{t=i}^{st=1,\dots,5}, WD_{t=i}^{st=1,\dots,5}, TP_{t=i}^{p=1,2}, CP_{t=i}, MWS_{t=i}) \\ \eta_{t=i+5} &= f(SSL_{t=i}, SLP_{t=i}^{st=1,\dots,5}, DSLP_{t=i}^{st=1,\dots,5}, WS_{t=i}^{st=1,\dots,5}, WD_{t=i}^{st=1,\dots,5}, TP_{t=i}^{st=1,2}, CP_{t=i}, MWS_{t=i}) \\ \eta_{t=i+12} &= f(SSL_{t=i}, SLP_{t=i}^{st=1,\dots,5}, DSLP_{t=i}^{st=1,\dots,5}, WS_{t=i}^{st=1,\dots,5}, WD_{t=i}^{st=1,\dots,5}, TP_{t=i}^{st=1,2}, CP_{t=i}, MWS_{t=i}) \\ \eta_{t=i+24} &= f(SSL_{t=i}, SLP_{t=i}^{st=1,\dots,5}, DSLP_{t=i}^{st=1,\dots,5}, WS_{t=i}^{st=1,\dots,5}, WD_{t=i}^{st=1,\dots,5}, TP_{t=i}^{st=1,2}, CP_{t=i}, MWS_{t=i})\end{aligned}$$

図 3: ガウス正規方程式

t は時間, s_t は観測所の数, p は経度と緯度, η は GMDH ベースのモデルによって予測される高潮。

誤差

次に相関係数 (CC), 正規化二乗平均平方根誤差 (NRMSE), 標準偏差 (NRMSE)

$$NRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^{i=n} (\eta_{obs,i} - \eta_{fore,i})^2}}{(\eta_{obs,max} - \eta_{obs,min})}$$

$$CC = \frac{\sum_{i=1}^{i=n} (\eta_{obs,i} - \bar{\eta}_{obs})(\eta_{fore,i} - \bar{\eta}_{fore})}{\sqrt{\sum_{i=1}^{i=n} (\eta_{obs,i} - \bar{\eta}_{obs})^2} \sqrt{\sum_{i=1}^{i=n} (\eta_{fore,i} - \bar{\eta}_{fore})^2}}$$

$$STD = \sqrt{\frac{1}{n-1} \sum_{i=1}^{i=n} (\eta_{obs,i} - \bar{\eta}_{fore})^2}$$

図 4: 誤差

η_{obs} は観測されたサージ水位, η_{fore} は予測されたサージ水位

性能指数

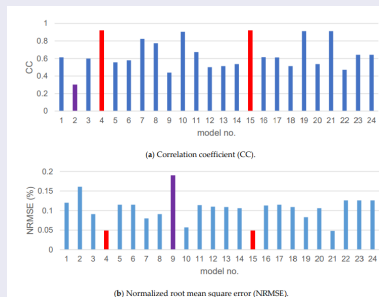


図 5: 性能指数

モデル 4 は再解析に最適なモデルであり．学習手順における最小残差 0.014 で 3 層から得られた．

$$y_3 = 0.00776 + 0.362533x_{13} + 0.598088x_{17_3} - 39.8235x_{13}^2 - 40.04x_{17_3}^2 + 79.9204x_{13}x_{17_3} \quad (13)$$

性能評価

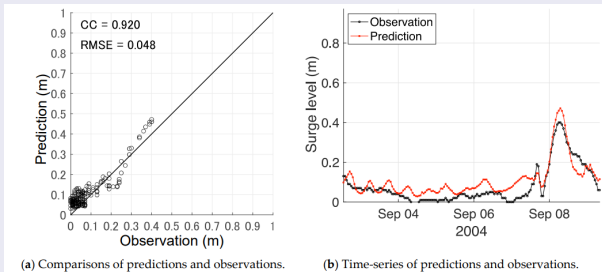


図 6: 性能指数

5時間リードタイム高潮予測モデル(サージ水位あり)

13/18

性能指数

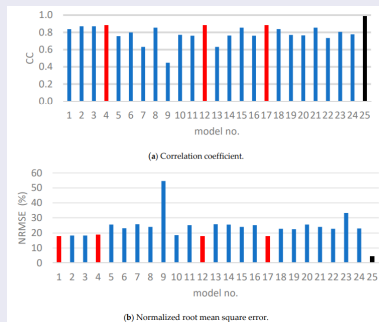


図 7: 性能指数

モデル 4 は再解析に最適なモデルであり、学習手順における最小残差 0.014 で 3 層から得られた。

$$y_3 = 0.00533 + 2.34286x_{4_4} + 1.42545x_{7_4} - 106.419x_{4_4}^2 - 99.754x_{7_4}^2 + 206.024x_{4_4}x_{7_4} \quad (14)$$

はじめに

GMDH

データ

テスト

結果

まとめ

性能評価

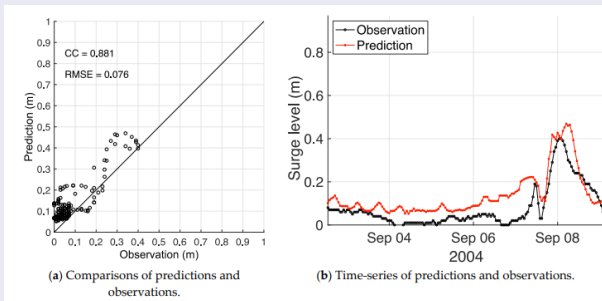


図 8: 性能指数

5時間リードタイム高潮予測モデル(サージ水位なし)

15/18

性能評価

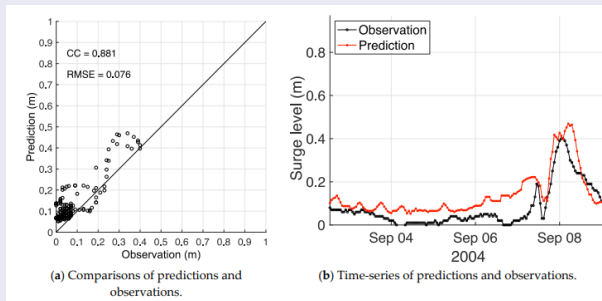


図 9: 性能指数

性能評価

同様に、『5 時間リードタイム高潮予測モデル (サージ水位あり)』

『12 時間リードタイム高潮予測モデル (サージ水位なし)』

『12 時間リードタイム高潮予測モデル (サージ水位あり)』

『24 時間リードタイム高潮予測モデル (サージ水位なし)』

『24 時間リードタイム高潮予測モデル (サージ水位あり)』

についても分析する。

性能評価

『5 時間リードタイム高潮予測モデル』

最も良い結果を示した訓練データセットは、高潮レベル、海面気圧、降水量、海面レベル、台風の位置、中心気圧の組み合わせ

『12 時間リードタイム高潮予測モデル』

高潮レベル、風速、風向

『24 時間リードタイム高潮予測モデル』

すべての入力パラメータが含まれる

現在の手法では ANN や GP の方法と比較して精度が低かった

現在の手法ではサージ水位の有無によって精度に差があり、5 時間および 12 時間では精度が向上。24 時間では同じ精度となった

適切な訓練セットとデータセットによって、現在の GMDH 手法の精度を向上させることが可能であることが分かった

今後の展開

気候変動を考慮した大規模なデータセットを用いて、GMDH ベースの高潮予測モデルの精度を向上させるための研究が必要。

アンサンブル予測システムを用いた GMDH ベースの高潮予測モデルの不確実性を実用的な予測に向けて検証する必要がある。

要素を自動的に決定する選択手法について、他の人工データセットとの比較を検討する必要がある。