

# 気圧配置分類に有効な類似度計算法の調査

\*佐藤拓人 (筑波大学・生命環境科学研究科), 日下博幸 (筑波大学・計算科学)

## 1. はじめに

気象予報や気象現象の解析の際、気圧配置図を分類する機会は数多くある。データの大規模化を考慮すると、分類手法の中でも、計算機を用いた客観的な分類が今後ますます注目されるだろう。

客観的分類のための類似度として相関係数や二乗誤差、ユークリッド距離が類似度として使われてきた。最近では, Mo et al. (2014) が、気象分野に限らず相関係数に関連した類似度計算法を実装し、相互比較を行った。しかし、画像ハッシュのように、相関係数とはコンセプトの異なる類似度計算法に関する検討は未だ不十分である。また、単一の典型的な教師データを設定した実験しか行っていない点も問題であろう。

そこで本研究では、5 つの類似度計算法(ユークリッド距離、相関係数、structural similarity, average hash, S1-score)を実装し、複数の教師データを設定して精度を相互比較する。

## 2. データと手法

使用するデータは JRA-55(Kobayashi et al, 2016)の日本周辺の 10 年分の海面更正気圧である(14,612 枚)。対象とした気圧配置型は日本海低気圧型である。本研究では、複数の専門家(N=5)による目視で、80%以上(N=4以上)が日本海低気圧と判断したものを日本海低気圧(328 枚)とした。また同様の方法で非日本海低気圧データ(788 枚)も作成し、調査用のデータセットとした。

日本海低気圧型のうち 1 つを教師データとし、その他のデータセット内の全てのデータとの類似度を計算する。類似度の高い順にデータセットを並べ替え、先頭から特定の数データを抽出し、取り出すことができた日本海低気圧型の全体に対する割合を抽出率とした。特に全ての日本海低気圧型(328 枚)を教師データとして抽出率の評価を行うことで、教師データによってどの程度差があるかを調査した。

## 3. 結果

表 1 に、各類似度計算法における精度を示す。Structural similarity (SSIM)と S1-score (S1)は、全ての教師データで評価した場合の平均的な抽出率が、他の類似度計算法より 10%以上高い。そのため、教師データ

によらず、他の手法より高精度になりやすいと言える。また、全ての教師データで評価した場合に最低となった抽出率を見ても、SSIM と S1 はその他の手法より高く、教師によらずある程度の精度で抽出できることがわかった。また、抽出されたデータセットをみると、SSIM と S1 は教師データのノイズに対してある程度ロバストで、視覚的な類似性をよくとらえることができた例があった。これに対して、従来使用されていたユークリッド距離(EUC)は、低気圧の類似性はとらえられたが、視覚的な類似性はそれほどとらえられない場合もあった。

## 4. 結論

本研究で対象とした日本海低気圧型の場合、Structural similarity と S1-score が教師データによらず高精度で分類が可能である。また、類似度によって重視する要素に違いがあることが確認され、特に視覚的な類似性をよくとらえた類似度が高精度になったものと思われる。

表 1 各類似度計算法の最高・最低・平均の抽出率(%)。抽出率は、類似度の上位 328 位までに含まれる日本海低気圧型の枚数の割合とした。COR: 相関係数, EUC: ユークリッド距離, aHash: average hash, SSIM: structural similarity, S1: S1-score をそれぞれ表す。

Method	Maximum	Minimum	Mean
COR	73.1	19.6	43.2
EUC	64.2	20.2	39.2
aHash	66.7	15.6	40.6
SSIM	77.7	26.9	56.7
S1	80.4	23.9	57.7

## 謝辞

本研究の一部は、(独) 環境再生保全機構の環境研究総合推進費 (2-1905) により実施された。

## 参考文献

- [1] Mo R. et al., 2014, *J. Hydrometeorol.*, **15**, 1862-1880.
- [2] Kobayashi C. et al., 2016, *J. geophys. Res. Atmos*, **121**, 1493-1510.