



九州大学



2018年12月19日

国立研究開発法人海洋研究開発機構

国立大学法人九州大学

国立研究開発法人科学技術振興機構

人工知能を用いて気候実験データから熱帯低気圧のタマゴを高精度に検出する
新手法を開発
～台風発生予測の高精度化に期待～

1. 概要

国立研究開発法人海洋研究開発機構（理事長 平 朝彦、以下「JAMSTEC」という。）地球情報基盤センターの松岡大祐技術研究員（国立研究開発法人科学技術振興機構さきがけ研究者）らは、国立大学法人九州大学大学院システム情報科学研究所の内田誠一主幹教授らと共同で、ディープラーニング（※1）によって、全球雲システム解像モデルNICAM（※2）による気候実験データから、発生前の熱帯低気圧の予兆を示す雲（熱帯低気圧のタマゴ、図1）を精度よく検出する手法（図2）を開発しました。開発した手法は特に夏の北西太平洋において、発生1週間前の熱帯低気圧のタマゴを高精度に検出可能であることを示しました（図3、4、5）。本成果により、人工知能（AI）技術を活用した新しい台風発生予測の実現に向けて大きな手掛かりが得られたと言えます。

本研究は、これまでの物理方程式に基づく気象モデルを用いた Model-driven（※3）な手法による将来予測の課題を克服すべく、過去に蓄積された大量のシミュレーションデータから現象発生の予兆を示す特徴を直接的に学習し、熱帯低気圧の発生を予測しようとする新たなアプローチの研究と考えられます。大量の気象ビッグデータが蓄積されていく現在において、本研究の成果は Data-driven（※4）な手法を用いた気象予測の新たな展開を拓くものとして期待されます。

本研究は科学技術振興機構 戦略的創造研究推進事業（さきがけ）の支援を受け、情報計測（計測技術と高度情報処理の融合によるインテリジェント計測・解析手法の開発と応用）領域の平成29年度採択課題「気象ビッグデータからの極端現象発生予測～台風のタマゴ発見から豪雨予測まで～」において行われたものです。また、本研究の一部は、JSPS 科研費（JP16K13885、JP26700010、JP17K13010）の助成を受けて行われました。

本成果は、日本地球惑星科学連合の英文論文誌「Progress in Earth and Planetary Science」電子版に12月19日付け（日本時間）で掲載されました。

タイトル：Deep Learning Approach for Detecting Tropical Cyclones and their Precursors in the Simulation by a Cloud Resolving Global Nonhydrostatic Atmospheric Model

著者：松岡大祐^{1,2}、中野満寿男³、杉山大祐¹、内田誠一⁴

1：JAMSTEC 地球情報基盤センター

2：科学技術振興機構戦略的創造研究推進事業（さきがけ）

3：JAMSTEC シームレス環境予測研究分野

4：九州大学大学院システム情報科学研究院

2. 背景

台風やハリケーン等のような熱帯低気圧の発生予測は、衛星観測による雲の発達具合の監視や、観測データを入力した気象モデルによるシミュレーションによって行われています。しかし、一般に大気現象は非線形性が強いいため、気象モデルによって遠い将来を予測しようとすればするほど予測結果のばらつきは大きくなります。

他方、近年の AI 技術の発展は目覚ましく、ビッグデータ中に潜む特定のパターンを反復的に機械学習し特定の事象を検出するディープラーニングが、さまざまな分野（機械音声認識、車両の障害物センサー、癌の検出、マーケティングの自動化等）でイノベーションを起こしており、不確実性の高い気象分野においても今後の応用が期待されています。

そこで JAMSTEC の松岡大祐技術研究員らは、画像認識の専門家である九州大学大学院システム情報科学研究院の内田誠一主幹教授と研究グループを構成し、ディープラーニングを用いて大量のシミュレーションデータから熱帯低気圧発生の予兆となる状態（タマゴ）を検出するという Data-driven なアプローチを試み、その検出精度を検討しました。

ディープラーニングを用いて優れた識別精度を得るためには、1 種類あたり数千枚を超える膨大な量のデータが必要となります。本研究では、まず、NICAM による 20 年分の気候実験データに対して熱帯低気圧の追跡アルゴリズムを適用することで、熱帯低気圧のタマゴおよび発達中の熱帯低気圧の雲画像 5 万枚を生成しました。さらに、熱帯低気圧に発達しなかった低気圧の雲画像 100 万枚と合わせて計 105 万枚の画像から学習データのセットを 10 組作成し（図 1）、深層畳み込みニューラルネットワーク（※5）を用いた機械学習によって、特徴の異なる計 10 種類の識別器を生成しました。そして、10 種類の識別器による結果を総合的に評価することで最終的な判断を行うアンサンブル識別器を構築しました（図 2）。

3. 成果

構築したアンサンブル識別器を、未学習のデータ 10 年分に適用し、熱帯低気圧およ

びタマゴがどの程度正しく検出できるか検証を行いました。実験結果の評価指標には、対象とするデータ中に存在する熱帯低気圧またはタマゴのうち、どの程度を正しく検出できたかを表す捕捉率と、熱帯低気圧・タマゴと予測した結果のうち、どの程度が間違えていたかを表す空振り率を用いました。その結果、最も良い検出精度が得られた例の一つでは、9個の熱帯低気圧・タマゴのうち8個を正しく検出することに成功しつつ（捕捉率88.9%）、熱帯低気圧・タマゴだと予測した82領域のうち、誤検出であったのはわずか8領域（空振り率9.8%）でした（図3）。ここで、捕捉率は対象とするデータ中に存在する熱帯低気圧およびタマゴのうち、どの程度を正しく検出できたかを表し、空振り率は熱帯低気圧またはタマゴであると予測した結果のうち、どの程度が間違えていたかを表します。

また、海域ごとに検出性能を評価したところ、平均的には北西太平洋が最も検出性能が高く、北インド洋が最も低いという結果が得られましたが、海域ごとの検出性能は各海域における熱帯低気圧の寿命の長さや学習データの数に強く依存することが明らかになりました。特に北西太平洋の台風シーズン（7月～11月）の熱帯低気圧・タマゴについては、捕捉率79.0%～89.1%という高い数値であるのに対し、空振り率は32.8%～53.4%と比較的低く、高い検出性能を得られることが分かりました（図4）。熱帯低気圧のタマゴは時刻を遡るにつれて捕捉率は低くなりますが、北西太平洋では10年間の平均で、発生2日前、5日前、7日前のタマゴのうち91.2%、77.8%、74.8%を検出することに成功しました（図5）。

4. 今後の展望

本成果によって、NICAMによる気候実験データを用いた熱帯低気圧のタマゴの検出に限っては、高い検出性能が得られました。一方で、現実の熱帯低気圧の発生を事前に予測するためには、データ同化を行ったシミュレーションデータや、衛星観測によって得られた雲画像に対しても同程度以上の検出性能が得られるよう、最先端の情報科学または統計数理的な手法を取り入れ、引き続き検討を進める予定です。

本研究において用いたディープラーニングは、画像認識だけでなく株価や乗客数の将来予測といった時系列変化に関する研究も盛んに進んでおり、台風の経路予測や強度予測、集中豪雨の発生予測等にも応用することが可能であると考えられます。今後、ディープラーニングに代表される人工知能技術を、他の応用分野からいち早く取り入れることによって、Data-drivenな手法とModel-drivenな手法を融合させた新しい海洋地球ビッグデータ解析の展開が期待されます。

[用語解説]

※1 ディープラーニング：多層化したニューラルネットワーク（ディープニューラルネットワーク）を用いた機械学習の一手法であり、画像認識、音声認識、言語理解、行動認識等のさまざまな分野において従来の性能を圧倒する大きなブレイクスルーを起こしている。近年のインターネットやコンピュータ技術の普及による膨大な量のデータの蓄積に加えて、複雑なタスクを可能とするニューラルネットワークの多層化技術の革新、それらの処理の高速化を実現するための計算機技術（特にGPU）の発展がベースとなり、2012年以降、第3次人工知能ブームの火付け役として、近年、大きな注目を集めている。

※2 NICAM：地球全体で雲の発生・挙動を直接計算することにより高精度の計算を実現した全球気象モデル。従来の全球気象モデルでは、高気圧・低気圧のような大規模な大気循環と雲システムの関係について、なんらかの仮定が必要とされ、不確実性の大きな要因となっていた。NICAMは主に水平解像度870mから14kmの範囲で運用されており、870m～3.5kmの超高解像度を用いる場合は全球雲解像モデル、7km～14kmの解像度を用いる場合は全球雲システム解像モデルと呼ばれる。台風の再現性能が高く、これまでに台風の研究で数々の成果をあげている。本研究において使用した気候実験データは、文部科学省のHPCI戦略プログラム分野3「防災・減災に資する地球変動予測」のもと、理化学研究所のスーパーコンピュータ「京」を利用して得られたものである（課題番号：hp120279, hp130010, hp140219）。

(参考)

「地球温暖化が台風の活動と構造に及ぼす影響—強風域拡大の可能性を示唆—」（2017年9月14日プレスリリース）

ス) https://www.jamstec.go.jp/j/about/press_release/20170914/

「全球高解像度シミュレーションにより台風予測精度が向上—複数のモデルと多数の事例で高解像度化の効果を定量的に確認—」（2017年3月31日プレスリリース）

ス) https://www.jamstec.go.jp/j/about/press_release/20170331/

「台風発生の2週間予測が実現可能であることを実証—台風発生予測の実用化に向けた第一歩—」（2015年1月20日プレスリリース）

ス) https://www.jamstec.go.jp/j/about/press_release/20150120/

※3 Model-driven（モデル駆動）：理論に基づいて自然現象をモデル化した物理モデルを用い、演繹的に解を導こうとする手法を指す。物理方程式に基づくシミュレーションモデルであるNICAM等が該当し、気象学等の自然科学分野では主流の研究手法として用いられてきた。

※4 Data-driven (データ駆動) : Model-driven とは逆に、得られたデータに対して統計数理的な処理を施すことによって、現象を支配している法則や関係式を帰納的に導こうとする手法を指す。本研究では、Model-driven な手法によって得られたシミュレーションデータに対して、Data-driven な手法を適用している。

※5 畳み込みニューラルネットワーク : 特に画像認識に特化したディープニューラルネットワークのモデルの一つであり、画像中に写った顔や物体の認識、自動運転技術や工場生産における異常検知、CT スキャンデータからのがん細胞検出にも応用されている。従来の画像認識では、認識したい対象物を表す特徴量 (画像の濃淡パターン等) を人間が与える必要があったが、ディープラーニングでは、大量のデータを用いることで、特徴量そのものをデータから自動的に学習でき、それにより大幅に認識性能を向上させることが可能となった。

お問い合わせ先 :

(本研究について)

国立研究開発法人海洋研究開発機構

地球情報基盤センター 先端情報研究開発部 情報・計算デザイン研究開発グループ
技術研究員 松岡 大祐 TEL:045-778-5862 E-mail:daisuke@jamstec.go.jp

国立大学法人九州大学 大学院システム情報科学研究院

教授 (主幹教授) 内田誠一 TEL:092-802-3574 Email:uchida@ait.kyushu-u.ac.jp

(報道担当)

国立研究開発法人海洋研究開発機構

広報部 報道課長 野口 剛 TEL:046-867-9198 E-mail:press@jamstec.go.jp

国立大学法人九州大学

広報室 TEL:092-802-2130 E-mail:koho@jimu.kyushu-u.ac.jp

国立研究開発法人科学技術振興機構

広報課 TEL:03-5214-8404 E-mail:jstkoho@jst.go.jp

(JST 事業について)

国立研究開発法人科学技術振興機構

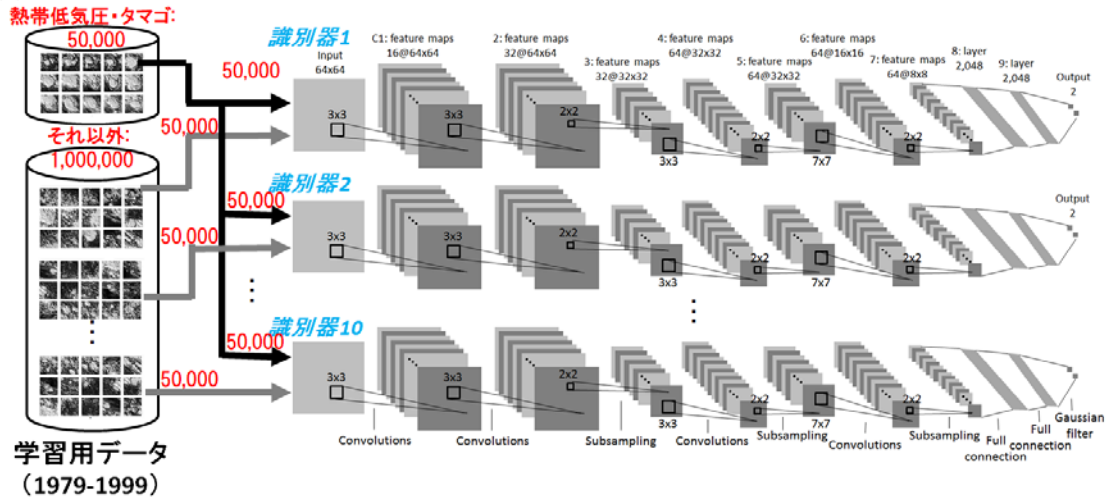
戦略研究推進部 グリーンイノベーショングループ

中村 幹 TEL:03-3512-3531 E-mail:presto@jst.go.jp



図 1. 熱帯低気圧およびタマゴの雲画像（外向き長波放射）の一例。NICAM 気候実験データ 20 年分に対して熱帯低気圧の追跡アルゴリズムを適用することで生成した。各画像は $1,000\text{km}^2$ (64×64 グリッド) とした。

(a) 学習フェーズ



(b) 予測フェーズ

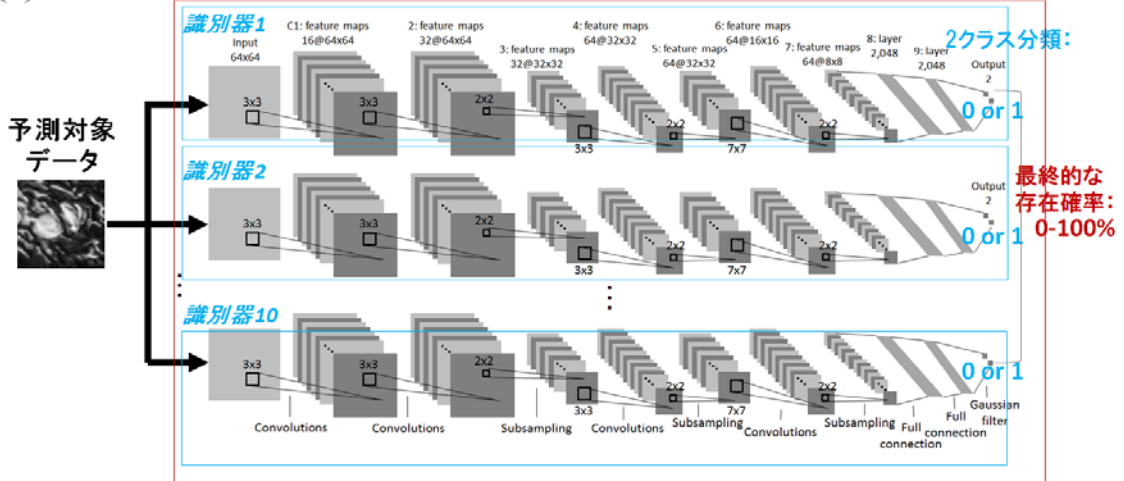


図2. 深層畳み込みニューラルネットワーク (CNN) によるアンサンブル識別器。(a) 学習フェーズでは、10種類のCNNがそれぞれ異なる学習データを用いて学習を行う。(b) 予測フェーズでは、1枚の入力画像に対して10種類の識別器を用いて2クラス分類を行い、それらの加重平均を最終的な存在確率とする。存在確率が事前に与えた閾値を超えた領域に対して、熱帯低気圧またはそのタマゴが検出されたとみなす。

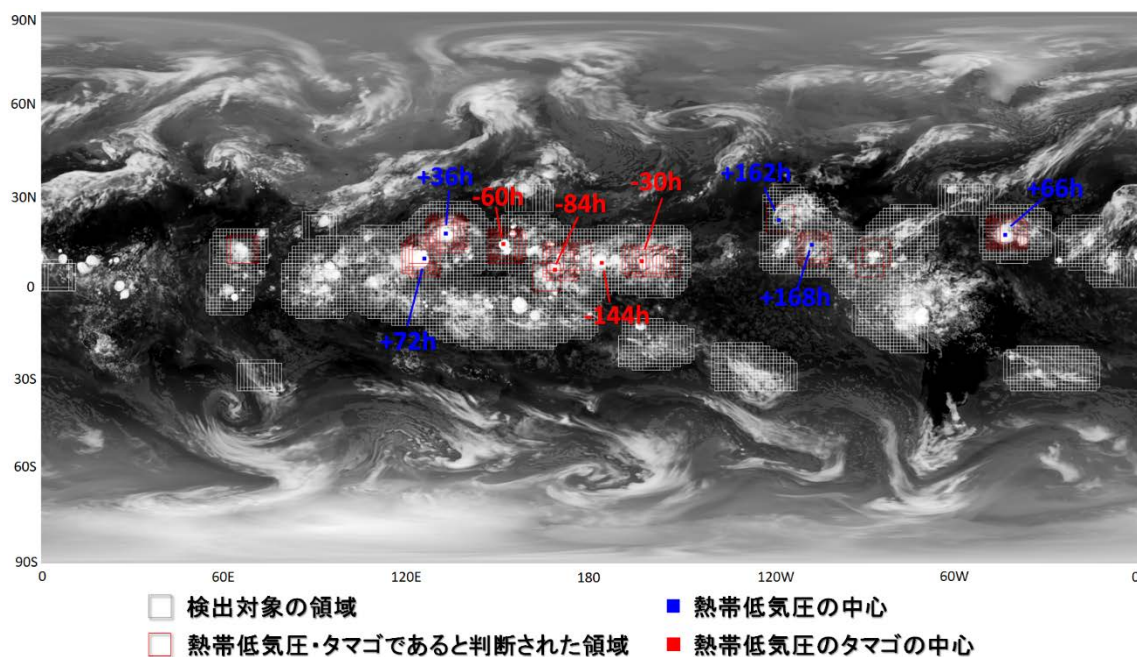


図 3. アンサンブル識別器を用いた未学習のデータに対する予測結果の一例。雲量 30%-95%の領域 (1,000km 四方) を予測対象とする領域とし白枠で示している。また、対象領域の雲画像に対して、熱帯低気圧またはタマゴの存在確率が 100%であると予測 (10 台の識別器全てが熱帯低気圧またはタマゴであると予測) した領域を赤枠で示している。青色および赤色の点は、熱帯低気圧の追跡アルゴリズムによってすでに分かっている熱帯低気圧およびタマゴの正解の中心点をそれぞれ表している。

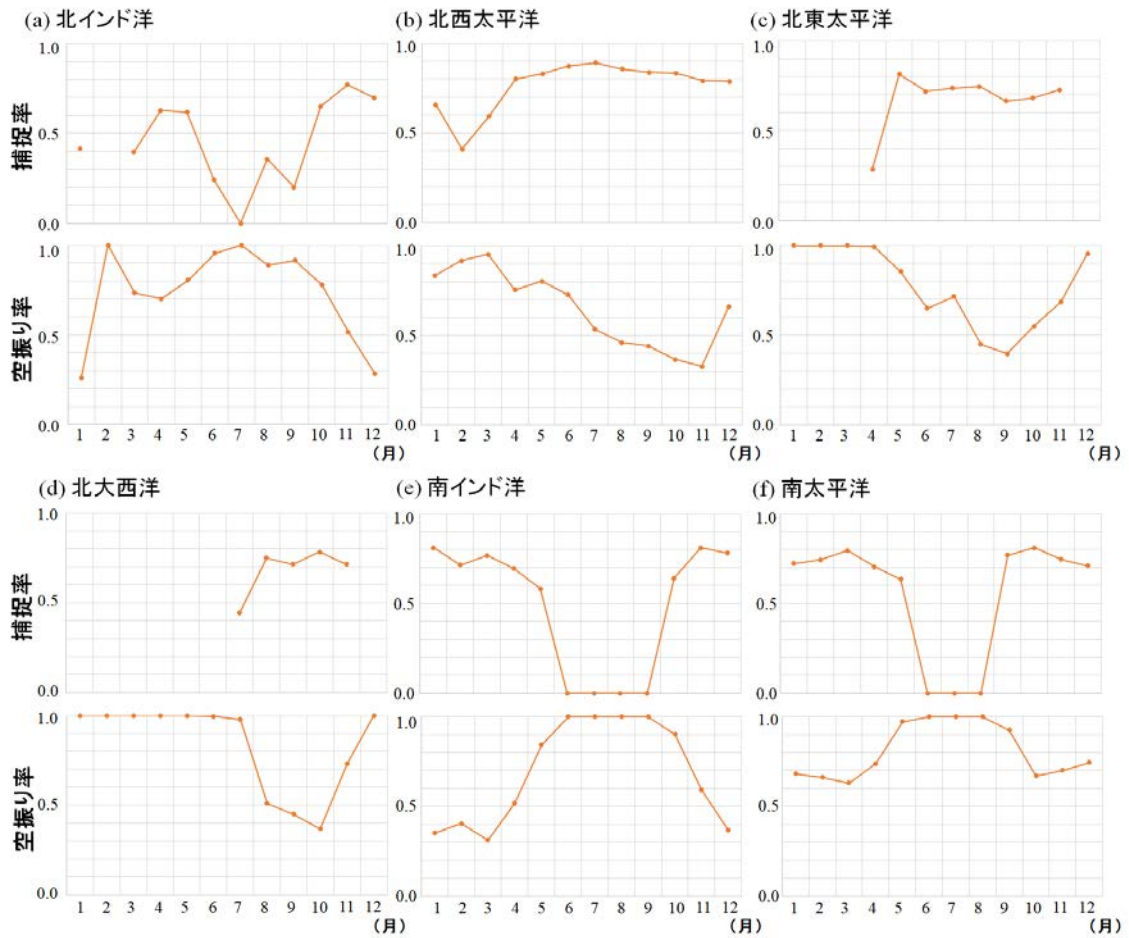


図 4. 海域毎の捕捉率および空振り率の月変化 (10 年平均値)。

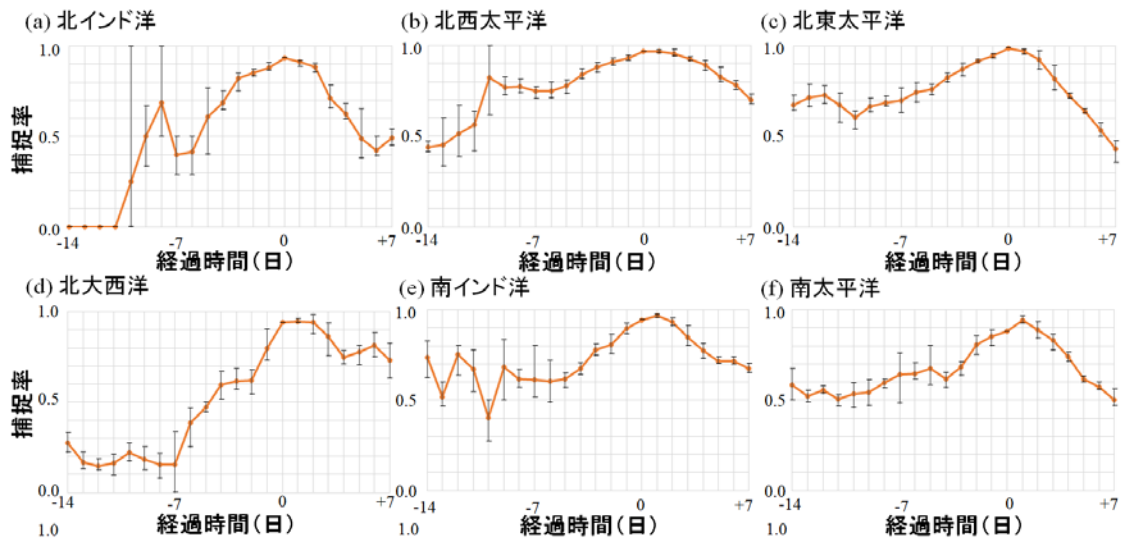


図 5. 海域毎の経過時間別捕捉率。経過時間の正の値は熱帯低気圧発生後の日数を表し、負の値は発生までの日数を表す。