

研究終了報告書

「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～」

研究期間：2017年10月～2021年3月

研究者：松岡 大祐

1. 研究のねらい

台風や豪雨のような極端気象現象に起因する水災害、土砂災害または風災害は、国民が安心・安全な社会生活を送る上で深刻な脅威となっている。特に近年の平均気温または海水温の上昇も一因となり、その極端現象そのものおよび災害の規模は大きくなっている。極端現象の発生を事前に予測することは、学術的に重要な課題であるだけでなく、社会的にも防災・減災の観点から重要な課題である。

近年、スーパーコンピュータの性能向上にともない、リアルタイム衛星観測データをデータ同化によって取り込んだ数値シミュレーションを用いた気象予測が主流となっている。一方で極端気象現象の予測には、非線形性にともなうシミュレーションの初期値依存性など、プロセス駆動型アプローチ特有の多くの不確実性がともなう。そこで本研究課題では、過去の観測データおよびシミュレーションデータから、機械学習を用いて現象発生の予兆となる時空間パターンのもつ特徴を抽出し、将来予測に活用するというデータ駆動型アプローチを採用する。

具体的な目標としては、台風などの熱帯低気圧発生の予兆の早期発見や、集中豪雨の高精度な将来予測を目指す。台風の予兆検出においては、過去のシミュレーションデータと観測データを融合した機械学習により、観測データではラベル付けが困難な発生初期の台風のタマゴデータを補間し、リードタイムの長い高精度な検出を実現する。また、豪雨予測においては、複数の初期値を用いたアンサンブルシミュレーションデータから、機械学習を用いて観測値に近いアンサンブルメンバーとそうでないアンサンブルメンバーとを識別し、より不確実性の小さい将来予測を実現する。

さらに、一連の研究開発において得られた地球科学ビッグデータに対する機械学習の活用に関する技術および知見を、領域内での共同研究を通じて、大気科学や水産学等の領域に適用する。

2. 研究成果

(1) 概要

本研究では、データ駆動型アプローチ(ビッグデータと機械学習)とプロセス駆動型アプローチ(気象モデルによるシミュレーション)の融合による気象予測の高度化を目指し、熱帯低気圧および豪雨予測を対象とした研究開発を行った。熱帯低気圧の予測については、過去30年分の気候シミュレーションデータを用い、深層畳み込みニューラルネットワークを用いた機械学習によって熱帯低気圧発生の予兆となる雲分布の特徴量を抽出し、その早期検出を可能にする技術の構築に成功した(研究テーマA)。これらの研究を雲画像データにおける不均衡な2クラス分類問題に置き換え、多数派クラスのデータ選別および少数派クラスのデータ拡張、重み付き

誤差関数などに基づく高精度なクラス分類手法を開発した(研究テーマ B)。また、観測データに基づく大気再解析データおよび現業機関において予報官が手動作成した天気図データを用い、降雨の原因となる停滞前線を機械学習によって自動的に抽出する手法の開発に成功した(研究テーマC)。開発手法による結果はシミュレーションデータ中の停滞前線の抽出にも適用された。さらに、気象予測を行うシミュレーションにおいて大気重力波の影響を含めるために用いられるパラメタリゼーションについて、背景風を考慮できない1次元鉛直近似計算という課題を克服するため、機械学習に基づく2次元的なパラメータ抽出手法の開発に成功した(研究テーマD)。得られた一連の成果は、査読付き原著論文3件、書籍1件として発表され、Progress in Earth and Planetary science 誌の The Most Downloaded Paper Award 2020 の受賞やアメリカ地球物理学連合(AGU)の Research Spotlight への選出など、国際的に高く評価された。

(2) 詳細

研究テーマ A「気象シミュレーションデータを用いた熱帯低気圧の予兆検出」

深層畳み込みニューラルネットワークを用いた機械学習を用い、20年分の高解像度気象シミュレーションデータ(NICAM、水平解像度14km)から、発生前の熱帯低気圧のタマゴを表す特徴的な雲パターンの学習を行った(図1)。提案手法は、5万枚の正例(熱帯低気圧および予兆)画像および、50万枚の負例(それ以外)画像を識別する2クラスの分類学習器から構成されている。それぞれ異なる10種類のデータを用いて学習を行った分類器を組み合わせたアンサンブル学習器により、高精度な検出を実現している。学習済みのディープニューラルネットワークを未学習のデータ10年分に適用した結果、特に北西太平洋における7月から11月の台風は、約80~90%程度の捕捉率かつ約33~53%の空振り率で検出を行うことが可能であるという結果が得られた。また、台風発生2日前、5日前、および7日前のタマゴのうち、それぞれ91.2%、77.8%、および74.8%を正しく検出することに成功した(Matsuoka et al., PEPS, 2018)。本成果は、Progress in Earth and Planetary science 誌の Most Downloaded Paper Award 2020 を受賞するなど、国際的に高く評価された。

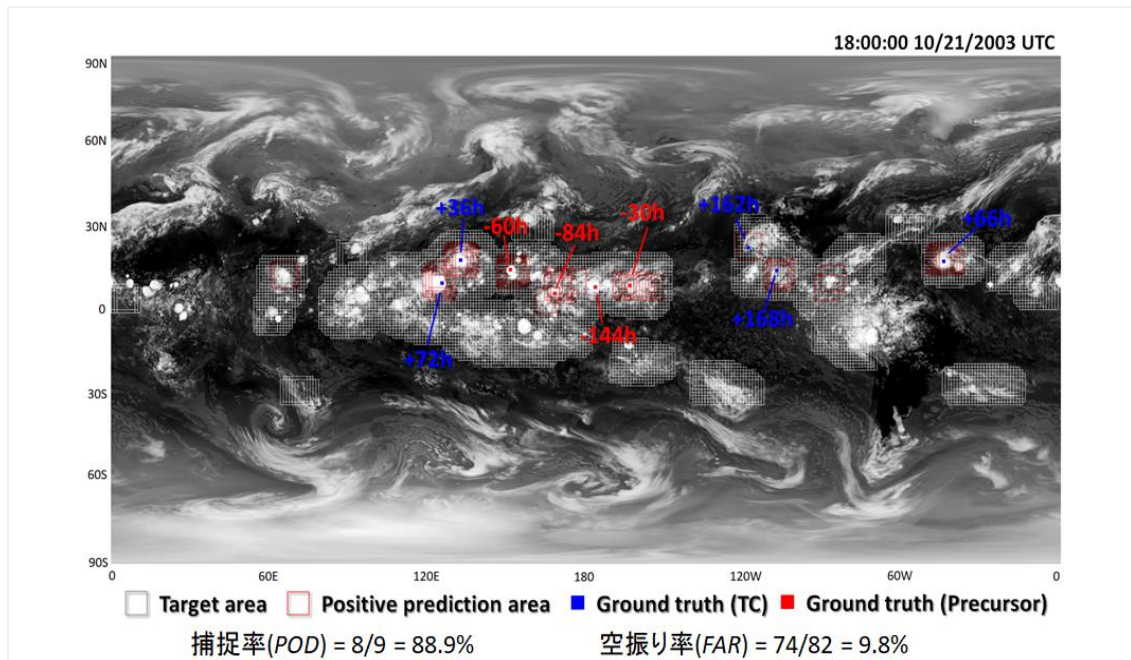


図 1 熱帯低気圧の予兆検出例

研究テーマ B「気象データにおける不均衡データの 2 クラス分類」

前述の熱帯低気圧の予兆の検出精度のさらなる向上を目指すため、不均衡データの高精度な 2 クラス分類に関する研究開発を行った。研究テーマ A では、2 クラスの画像数の不均衡さが、分類精度のさらなる向上に向けての課題となっていた。本研究では、気象データにおける不均衡データの 2 クラス分類について、①多数派クラス(負例)の選抜、②少数派クラス(正例)の水増し、③誤差関数の再定義、④先端的なニューラルネットワークアーキテクチャの利用、の観点から、精度向上に関する分析を行った。

200 種類以上の異なる設定による実験の結果、図 2 に示すように、捕捉率(再現率)を Matsuoka et al., 2018 と同様に 80%程度に固定した場合に、空振り率の平均値を 60%から 30%程度にまで軽減させることに成功した。多数派クラスの選抜においては、訓練データを用いた学習時において、分類が容易な負例(Easy negative)と分類が困難な負例(Hard negative)とに分け、Easy negative の数を減らし、Hard negative の割合を一定以上含めることによって、効果が得られた。少数派クラスの水増しにおいては、画像の拡大・縮小、クロッピング、回転によって 1 枚の画像を複数枚に増やすことで、分類精度の向上が見られた。ニューラルネットワークによる出力値と真値との誤差を定義する誤差関数について、少数派である正例に対する推定誤差の重みを大きくすることで、正例の分類精度が向上した。また、ResNet や PyramidNet 等の State-of-the-Art なアーキテクチャの有効性が確認された(松岡, 2021)。

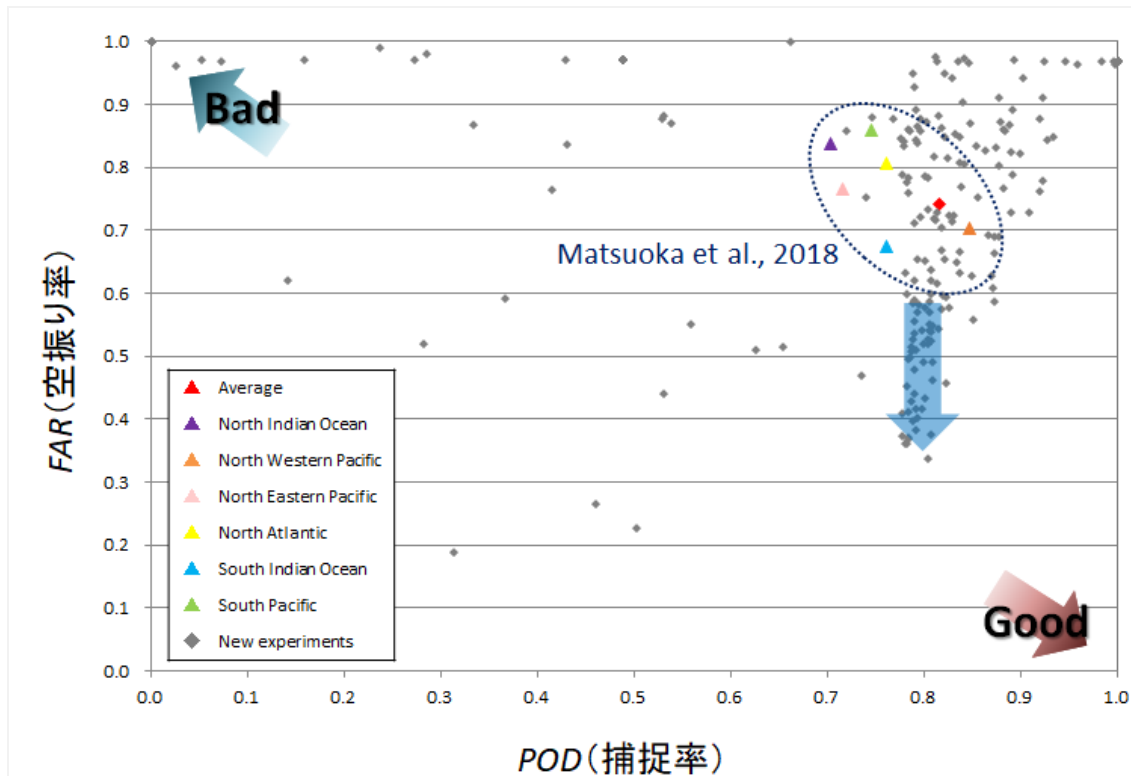


図2 熱帯低気圧の予兆検出の精度向上

研究テーマ C「停滞前線抽出による豪雨時の総観場の特定」

梅雨前線や秋雨前線のような季節性の停滞前線は、長期間の雨をもたらす河川の氾濫や浸水害、土砂災害等を引き起こすことがある。停滞前線は暖かい空気と冷たい空気の勢力が等しく釣り合った境界面として物理的に定義されるが、その位置を人間が直感的に解釈するには天気図等でもなじみのある停滞前線が分かりやすい。一方で天気図上の前線は、気象業務の担当者がデータを目視で確認して手動で描画するため、例えば将来の中長期の気象予測結果に対して前線の位置を描画することがコスト面で容易ではない。本研究では、気象シミュレーションによる将来予測データへの適用を目的とし、過去の観測データをベースとした再解析データを用い、機械学習(畳み込みニューラルネットワーク)による停滞前線の位置推定手法を開発した(図3左上および左下)。その結果、図3右に示すように、現業機関で用いられている地上の風や850hPa面の相当温位だけでなく、温度勾配や風向、水蒸気等を用いても前線の位置を精度良く推定することに成功した(Matsuoka et al., SOLA, 2019)。開発手法を過去実験および地球温暖化実験(4°C昇温時)のシミュレーションデータに適用し、地球温暖化気候実験データベース SEAL への登録および公開を行った。

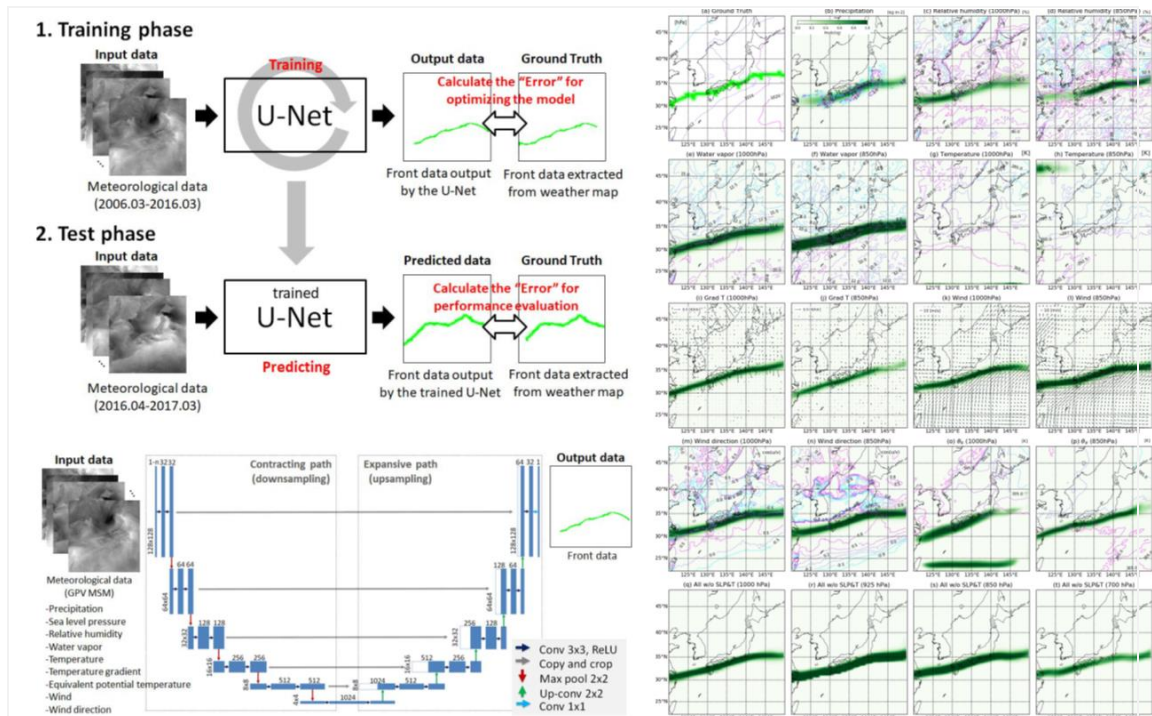


図3 深層畳み込みニューラルネットワークを用いた停滞前線の位置推定。(左上)学習、テストの流れ、(左下)ネットワークの構成、(右)真値および推定結果。

研究テーマD「機械学習を用いた大気重力波のパラメータ推定」

CREST 佐藤薫教授との共同研究において、機械学習を用いた大気重力波のパラメータ推定に関する研究開発を行った。大気重力波の空間スケールは、一般的な気象モデルの空間解像度と比べて小さいため、重力波の作用をパラメータ化して取り込む必要があるが、そのほとんどは鉛直方向の伝搬のみを仮定しており、背景風等による斜め方向の伝搬が考慮されていないという問題があった。

本研究では、深層畳み込みニューラルネットワークを用い、対流圏中層における低解像度(60km)の東西・南北風、温度、比湿データ等から、下部成層圏における高解像度(5km)の重力波フラックスの分布を推定する手法を開発した。開発手法を用いることで、特に山岳を風が乗り越える際に発生する地形性重力波による細かいスケールの重力波フラックスの分布を精度よく推定することに成功した(図4)。得られた成果は、アメリカ地球物理学連合(AGU)の発行するジャーナル 22 誌の全出版論文から毎月 10 数件のみが選出される Research Spotlight の一つとなるなど、国際的に高く評価された。

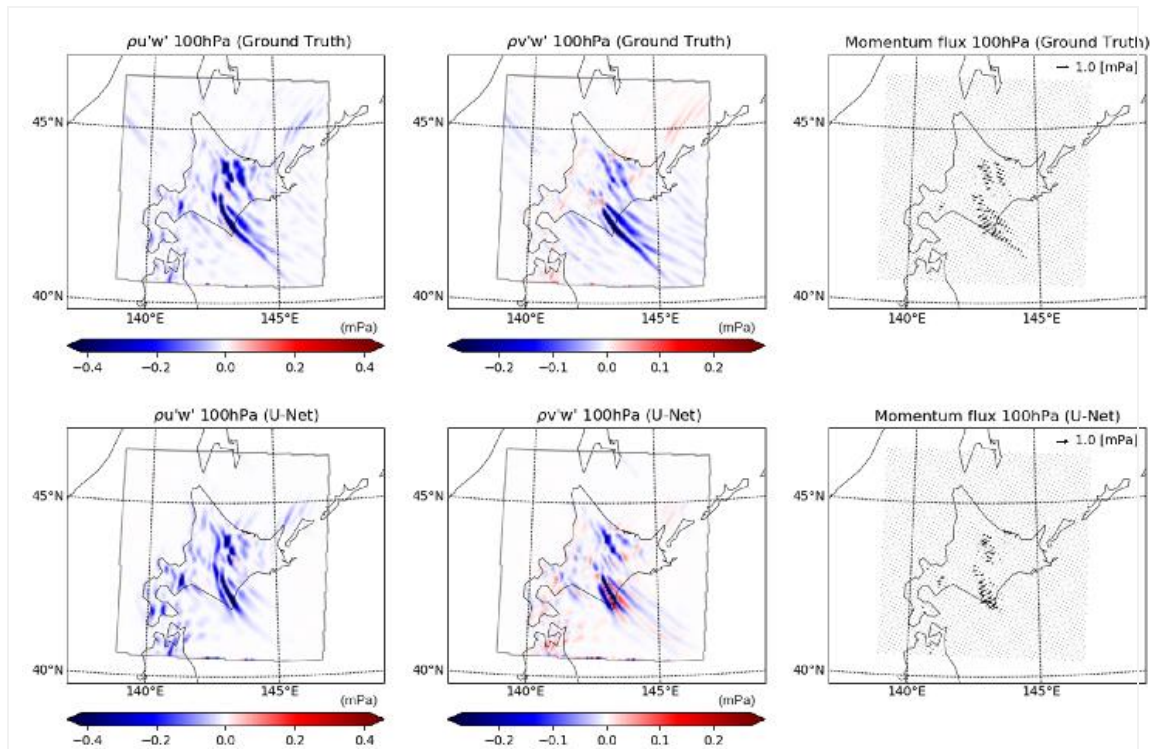


図 4 深層畳み込みニューラルネットワークを用いた運動量フラックスの推定値(下段)および真値(上段)。図左は東向き運動量の鉛直フラックス、中は北向き運動量の鉛直フラックス、右はそのベクトルを表している。

3. 今後の展開

一連の研究において、衛星観測データ、大気再解析データ、シミュレーションデータという異なる複数種のデータの融合、およびプロセス駆動型アプローチとデータ駆動型アプローチの融合において、特定の極端気象現象の予測の高度化や現象の理解につながる技術の基礎を構築した。や物理モデル化や観測が困難な、一市民スケール(個々の自宅や農場または漁場、船舶や自動車などの乗り物の周辺スケール)における、個々のニーズに即した予測への発展を検討している。その際、衛星観測データや再解析データ、シミュレーションデータだけでなく、個人所有のスマートフォンや、農場・漁場における定点カメラ、自動車や航空機などの車載カメラなどによる静止画・動画像などから、物理的な情報や環境的な情報を抽出し活用しようとする、次世代の気象予測への展開が考えられる。

4. 自己評価

研究開始当初に掲げた、シミュレーションデータと観測データの融合による気象予測の高度化という目標に対して、①シミュレーションデータによる学習と観測データへの適用、②観測データによる学習とシミュレーションデータへの適用、③観測データに基づく学習を用いたシミュレーションの高度化の3点において新たな技術を提案するに至った。現時点での論文掲載には至っていないが、④シミュレーションデータと観測データの融合による外挿的予測

を可能とする機械学習モデルの開発については、ベースとなる技術の開発は完了しており、台風の予兆検出においてその有効性を確認するに至っている。今後の気象分野以外への発展も期待できる。また、台風の経路予測や集中豪雨の予測については、目標を大きく定めすぎたこと、時間的な制約、計算機リソースの制約などの理由から、研究の完了には至らなかった。しかし、その基礎となる機械学習と大規模アンサンブルシミュレーションと観測データとを用いた予測の高度化について部分的には有効性が確認できており、研究期間終了後の速やかな論文投稿への見通しは立っている。

一連の研究成果は、PEPS (Progress in Earth and Planetary Science)誌の The Most Downloaded Paper Award 2020 を受賞するなど、国際的に高く評価された。また、自身が主導して行った領域内融合研究も、アメリカ地球物理学連合(AGU)の Research Spotlight として選出されるなど、国際的に高く評価された。ビッグデータとディープラーニングを活用した気象予測は、ここ数年の間に IBM や MicroSoft 等の超大手企業も参入し、非常に国際競争力の高い分野となっている。そのような中で、個人研究として社会・経済への波及効果を生むには現状では至っていないが、上述の理由により学術分野においては一定以上の存在感を示すことはできたと考える。本課題の成果の一部は、横浜国立大学およびアメリカ大気海洋庁(NOAA)が試験的に公開している台風発生環境場モニタリング Web サイトとして公開されているが、今後、気象庁などの現業機関との協働に発展させることによって、社会への波及効果が期待できる。

5. 主な研究成果リスト

(1) 代表的な論文(原著論文)発表

研究期間累積件数:3件

1. Daisuke Matsuoka, Masuo Nakano, Daisuke Sugiyama, and Seiichi Uchida, Deep learning approach for detecting tropical cyclones and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model, *Progress in Earth and Planetary Science*, 2020, Vol. 5, Issue 80, doi:10.1186/s40645-018-0245-y

複数の深層畳み込みニューラルネットワークからなる複数の 2 クラス分類モデルを組み合わせ、雲解像気象シミュレーション(NICAM)によるデータから、熱帯低気圧発生の予兆となる雲を高精度に検出することに成功した。PEPS The Most Downloaded Paper Award 2020 を受賞した。

2. Daisuke Matsuoka, Shiori Sugimoto, Yujin Nakagawa, Shintaro Kawahara, Fumiaki Araki, Yosuke Onoue, Masaaki Iiyama, Koji Koyamada, Automatic Detection of Stationary Fronts around Japan Using a Deep Convolutional Neural Network, *Scientific Online Letters for the Atmosphere*, 2019, Vol. 15, pages 154-159, doi:10.2151/sola.2019-028

深層畳み込みニューラルネットワークを用い、観測データに基づく大気再解析データから、降雨の原因となる停滞前線を自動的に抽出することに成功した。気象庁などの現業機関では、1000hPa 面の風や 850hPa 面の相当温位の分布を用いて手動で前線を描画していたが、本成果では水蒸気や温度勾配なども前線面を定義する上で重要であることを明らかにした。

3. Daisuke Matsuoka, Shingo Watanabe, Kaoru Sato, Sho Kawazoe, Wei Yu, and Steve Easterbrook, Application of Deep Learning to Estimate Atmospheric Gravity Wave Parameters in Reanalysis Data Sets, *Geophysical Research Letters*, 2020, Vol. 47, No. 19,

深層畳み込みニューラルネットワークを用い、下部成層圏の細かいスケールの大気重力波に関するパラメータ(運動量フラックスの最大振幅や卓越波数)を、対流圏中層の粗いスケールの風や温度や湿度から推定することに成功した。本論文はアメリカ地球物理学連合(AGU)の Research Spotlight に選出された。

(2)特許出願

研究期間累積件数:1 件(特許公開前)

(3)その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

学会発表(招待講演)

- ① Daisuke Matsuoka, Masuo Nakano, Daisuke Sugiyama, Seiichi Uchida, Deep Learning Approach for Detecting Precursors of Tropical Cyclone Simulated by a Global Nonhydrostatic Atmospheric Model, Japan Geoscience Union Meeting 2018, May 23, 2018
- ② Daisuke Matsuoka, Deep learning approach for detecting precursors of tropical cyclones -toward cyclogenesis prediction-, NOAA Emergency Technologies Workshop, June 25, 2019

受賞

- ① Daisuke Matsuoka, Masuo Nakano, Daisuke Sugiyama, Seiichi Uchida, PEPS (Progress in Earth and Planetary Science) The Most Downloaded Paper Award 2020 受賞, 2020年6月11日

著作物

- ① 松岡大祐, 杉山大祐, 畳み込みニューラルネットワークの気象データ解析への応用—熱帯低気圧の検出から雲量推定まで—, 可視化情報学会誌, Vol. 38, No. 151, Pages 133-137, 2018
- ② 松岡 大祐, 人と共生する AI 革命 ~活用事例からみる生活・産業・社会の未来展望~, 栗原 聡編, (株)エヌ・ティー・エス, 2019

プレスリリース

- ① 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～(2018年12月19日)
http://www.jamstec.go.jp/j/about/press_release/20181219/