

研究開発項目

2. データ駆動型気象予測手法の開発

2023 年度までの進捗状況

1. 概要

2023 年度中においては、データ駆動型の気象予測モデルの調査や予備実験に基づく基礎検討を行い、次年度以降に開発予定のプロトタイプおよびマイルストーンの詳細化を行いました。特に研究目的から想定される予測精度、学習データ量および利用可能な計算リソース（並列 GPU 計算機）から、Vision Transformer をベースとした代理モデルを開発することとしました。また、既存モデルをベースとして用いたいくつかの予備実験の結果から、当項目における開発の指針として、モデルの軽量化と予測精度のトレードオフを考慮可能な手法への拡張を検討することとしました。また、モデルの学習において必要なベンチマークデータの検討および準備を行い、複数の項目間で共有しました。

2. これまでの主な成果

2020 年以降に発表された 21 件の関連論文から、各モデルにおいて用いられている要素技術（機械学習手法、アーキテクチャ、前処理、誤差関数等）や使用データセット（物理量、時空間解像度等）、評価指標に加えて使用計算リソースや学習時間等を体系的に調査し、研究動向の把握に加えて分析を行いました。その結果、次年度以降に開発を進める代理モデルプロトタイプとして、表 1 に示すように、Swin Transformer 等のシンプルな構造をもつ Vision Transformer に加えて、同様に注意機構を有する畳み込みニューラルネットワークを基本的なアーキテクチャとし、大気再解析データの主要な 8 つの物理量 1 時間ステップ分のみを入力および出力値としてもつ自己回帰型のモデルとする方針が定まりました。

また、入力時の変数圧縮に関する検討と予備実験を実施しました。特に、既存モデル（ClimaX）のように全入力変数を 1 つの代表変数に圧縮するアプローチに対して、図 1 に示すように、明示的に複数の代表変数に圧縮する手法を提案しました。小規模モデルを用いた予備実験の結果、既存手法と比較してモデルのパラメータ数を約 50%削減しつつも予測精度を向上させることに成功しました。また、モデル開発に共通のベンチマークデータとして、現業気象予報において用いられるメソスケールモデル MSM に加え、

ダウンスケーリング版大気再解析データ DSJRA-55 によるデータセットの準備を行いました。

表 1: プロトタイプ開発における機械学習手法、要素技術および学習データ（一部抜粋）

Component	Detail
ML architecture	CNN (Attention R2U-Net), Vision Transformer (Swin Transformer)
Loss function	Latitude-Weighted MSE, Pressure-weighted loss, MAE, CRPS
Preprocessing	Z-score normalization for each 2D data
Input variables	3D: U, V, T, Q, RH, Z 2D: U10m, V10m, T2m, MSLP, Total precipitation (TP), OLR Fixed: Sea/land, Geopotential height
Output variables	Same as input (excluding fixed variables) -> Precipitation
Pressure level	13 level
Temporal resolution	1 h, 3 h, 6 h, 24 h
Forecasting approach	Direct forecasting, Continuous forecasting, Iterative forecasting
Training data	GPV-MSM, DSJRA55
Training approach	Base model + Fine-tuning (precipitation)

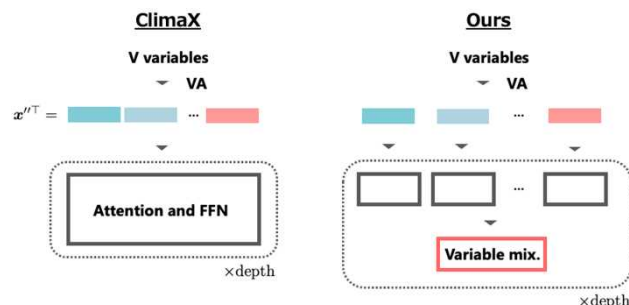


図 1: 既存モデル（ClimaX）と提案手法における入力変数の扱い方における違い

3. 今後の展開

予備検討および予備実験の結果に基づき、日本周辺域の豪雨現象の予測に特化した予測モデルの開発を継続します。特に、異なる時間解像度（例えば 1 時間、3 時間、6 時間、12 時間）の予測モデルの統合や、全球から領域スケールをシームレスにつなぐ予測モデルの開発を推進し、軽量かつ高精度な代理モデルの基盤技術を整備します。

また、共通ベンチマークデータに加えて、ベンチマークイベント（平成 26 年 8 月豪雨、平成 29 年 7 月九州北部豪雨、平成 30 年 7 月豪雨等）を設定し、代理モデルによる予測（項目 2）や分析（項目 3）、可視化（項目 10）等において協調して研究を進める予定です。