

AI気象モデル研究の進展と気象庁との連携 ——大学研究者からの視点——

本田匠

東京大学情報基盤センター

近年の動向（ごく一部）

- 現業機関
 - ECMWF(欧州): anemoiフレームワークの開発、AIFS Single/ENSの現業化
 - NOAA(米国): AIG(E)FSの現業化、領域版HRRRCastほか
 - ECCC(カナダ): AI予測場をnudgingし、物理モデルを駆動
- Big Tech
 - Google: Weather Next 3 (衛星観測を利用、1時間ごとに予測更新可)をリリース
 - NVIDIA: データ同化(HealDA)を含むEarth 2 familyをリリース
 - Microsoft: 基盤モデルAuroraの開発

AIにはできない（と思われていた）こと

- 高解像度・極端気象予測
 - [HRRRCast](#) (6-km mesh over CONUS)
- 確率/長期予測
 - AIFS ENSほか多数
- 物理的な整合性
 - 熱帯熱源からMatsuno–Gill likeなパターン([Hakim and Masanam 2024](#))
- データ同化
 - 観測+シミュレーションを入力: [Aardvark Weather](#)
 - 観測のみ入力: HealDA
 - 観測のみから予測まで([DOP](#), [OCELOT](#))

大学における研究の方向性

- Toy/軽量モデルを用いたproof of concept
 - 例: Lorenz 40変数モデルで理想化実験を構築必要なもの: **AI気象モデル構築フレームワーク**
- 既存のAI気象予測モデルの応用、拡張
 - 例: オープンソースモデルを取得、fine tuning、物理モデルと比較必要なもの: **オープンソースAI気象モデル、学習データ**
- AI気象予測モデルの構築
 - 例: オープンソースモデルのLossを改良、事前学習とfine tuning必要なもの: **オープンソースAI気象モデル、学習データ**

必要なモデル

- 重み(check point)
- ソースコード(モデルarchitecture)
- 実行手順(notebookなど)

AIFSやWeather Next (GraphCastなど)では（一部を除いて）公開
→（気象庁が今後開発するものが）要件を欠くor入手困難であれば、
おそらくAIFSなどを使うことになる

必要な学習データの要件

1. 高精度
2. 良アクセス
3. (既存モデルでの実績)

いずれもERA5がbest

ERA5ではなくて気象庁データを使うには、何かしらの理由づけが必要

- ERAと異なる予報・診断変数を含む
 - ERAを上回る解像度
 - ERA以上の良アクセス
- など？

雑多な所感1: AI気象モデルに必要な(再)解析データ

従来の気象・気候研究のニーズ $\neq$ AI気象モデル研究のニーズ

- 高精度○
- (水平)高解像度○
- (鉛直)高解像度△
 - 一部の指定気圧面だけでも十分
- 長期間△
 - >50年のデータがあっても、実際は~1980年以降しか使っていない
 - Fine tuning and/or 領域データであれば短期間(HRRRCastは~3年)でも十分?

雑多な所感2: AI気象モデルは未検証か？

- 基本はYes, 使用されてきた年月: AI気象モデル << 物理気象モデル
- ただし、（一般的な）研究ベースの物理気象モデルと比較すると差は小さい or むしろ検証が進んでいる？

トップAI気象モデルの開発・解説論文は多面的かつ実践的な検証が充実
(対解析, 対観測, 対物理気象モデル, 確率予測, 極端気象…)

後発研究の数も多い

おわりに

- 大学研究者も（少なくとも本田は）気象庁と連携し、気象庁のモデル・データを使いたい
- 大学側の事情
 - 時間や人員の不足 → ERA5 + AIFSなどを使う方が楽
 - GPU時間の確保（調達and/or公募）
 - 学術誌のオープンデータ/ソースの要求
- 大学研究者を含むユーザーと密なコミュニケーションを取りながら、AI気象モデルや再解析データほかの開発を進めていただければ幸いです