

物理法則に基づくニューラルネットワーク構築の検討 —対流圏上層の風予測を例に—

田中 潤也^{1,a)} 富田 智彦³ 沼尾 正行² 福井 健一²

概要: 近年、観測データからパターンを取得し、未知のデータに対して予測・分類を行う機械学習が多くの商業分野で成功している。しかし、原理を探究する自然科学分野では、予測や分類において高い精度が得られたとしても、モデルに説明性がなければ有用なモデルを獲得できたとは言えない。そこで自然科学分野で使用される物理モデルを機械学習モデルに統合することにより、このような機械学習の問題点を解決する手法に着目した。本稿では、対流圏上層の風を観測データから予測するタスクにおいて、温度風方程式を制約として畳み込みニューラルネットワークの損失関数に加え、物理モデルと機械学習モデルを統合し予測を行う手法について検討した。

キーワード: ニューラルネットワーク, 物理モデル, 温度風方程式, 地衡風

Theory-guided Learning of Neural Network — Prediction for Wind in the upper troposphere —

JUNYA TANAKA^{1,a)} TOMIHIKO TOMITA³ MASAYUKI NUMAO² KENICH FUKUI²

Abstract: In recent years, machine learning which obtains patterns from observation data and predicts and classifies unknown data has been successful in many commercial fields. However, in the field of natural sciences, even if high precision is obtained in prediction and classification, it can not be said that a useful model could be acquired without a model being descriptive. We focused on technique which combined with physical model used in natural science field and machine learning model to improve the explanation of the model while maintaining high accuracy. In this paper, we used an example to predict winds in the upper troposphere from temperature wind equations. We investigated whether the problem of machine learning can be solved by adding to the loss function of the neural network as constraints the temperature wind equation consisting of the horizontal components.

Keywords: Neural Network, Physics, Thermal Wind, Geostrophic wind

1. はじめに

近年、観測データからパターンを取得し、未知のデータに対して予測・分類を行う機械学習が多くの商業分野で使用されている。しかし、機械学習システムでは、ノイズ・

外れ値を含む観測データから学習しているため、既知の法則性を無視した出力を得る可能性がある [1]。また、ニューラルネットワークに基づく深層学習では、予測・分類は高精度で行える一方で、学習モデルが複雑になり人間がモデルを解釈することが困難になるという欠点があった [2]。特に、原理を探究する自然科学分野では、予測や分類において高い精度が得られたとしても、モデルに説明性がなければ有用なモデルを獲得できたとは言えない [3], [4]。

そこで、近年注目されている自然科学分野で使用される物理モデルを機械学習モデルに統合することにより、この

¹ 大阪大学 情報科学研究科 Graduate School of Information Science and Technology, Osaka University

² 大阪大学 産業科学研究所 The Institute of Science and Industrial Research, Osaka University

³ 熊本大学 大学院先端科学研究部 Faculty of Advanced Science and Technology, Kumamoto University

a) tanaka@ai.sanken.osaka-u.ac.jp

ような機械学習の問題点を解決する手法に着目した [5], [6]. 本稿では, 対流圏上層の風を温度風方程式から予測する例を使用した. 対流圏上層の風を観測データから予測するタスクにおいて, 東西成分と南北成分からなる温度風方程式を制約としてニューラルネットワークの損失関数に加え, 機械学習の問題点が解決できるかを検討した.

本稿の構成は次の通りである. まず, 2章で物理モデルと機械学習モデルを統合させることで, 双方の課題を解決している先行研究について述べる. そして3章で本稿で着目する手法等について説明する. 4章では, 実験内容と結果および考察について述べる. 最後に5章でまとめと今後の展望について述べる.

2. 先行研究

本章では, 自然科学分野で使用されている物理モデルと機械学習を統合させることによって, 双方の課題を解決している先行研究について述べる.

2.1 ドメイン知識に基づく教師あり学習

Stewart らは, 単純な運動方程式や既知のルールなどのドメイン知識を畳み込みニューラルネットワークの損失関数に制約条件として統合させることによって, 教師信号なしの観測データから教師あり学習を行う枠組みを提案した [7]. 深層学習などのニューラルネットワークを用いた機械学習モデルは, 高い精度で予測・分類を行える一方で, 大量の教師ありデータが必要であるという問題があった. 事前に観測データに潜むパターンがある程度把握できている場合に, ドメイン知識から制約を作り教師信号なしで学習できる枠組みを提案している.

この先行研究では, 教師あり学習の問題点である教師ありデータを使用せずに学習する方法に焦点が置かれ, また, 扱っている例も単純な運動方程式やルールなど, シンプルなものが多い.

2.2 畳み込みニューラルネットワークによるシミュレーションの高速化

気象学を含む流体を扱う分野では, 流体運動に関する Navier-Stokes 方程式の効率的なシミュレーションが長年にわたり課題となっている. Tompson らは, 深層学習の近似力を活用し, 高速で高精度なシミュレーション結果を獲得できるアプローチを提案した [8].

この先行研究では, 差分法と畳み込みニューラルネットワークを組み合わせ, 非圧縮流体シミュレーションを行っている. 畳み込みニューラルネットワークによる出力は通常非線形非凸最適化手法が適用される圧力算出部分であり, 入力には現在の格子点の速度, 位置, 圧力である. 煙のようなスパースな流体で良い精度を獲得している.

3. 物理モデルと機械学習モデルによる対流圏上層の風予測

Stewart らの研究では, 単純な運動方程式やルールでの有効性を検証している. 本研究は前章で Stewart らの先行研究をもとに, 自然科学分野のような複雑な問題に拡張し, 機械学習の問題点である既知の法則性を無視した出力を得てしまうといった問題を解決できるかを検討する. 先行研究よりさらに実問題に近い問題として, 温度風方程式を用いた対流圏上層の風を予測するタスクに着目した. 先行研究で挙げられていた物体の放物線運動より複雑な非線形現象であり, 微分方程式で近似することができること, またデータセットが容易に入手できるという観点から, 温度風方程式を使用した.

3.1 地衡風

地衡風とは, 気圧傾度力とコリオリ力の釣り合いの結果生じる風である [9]. 気圧差がある平面において, 気圧傾度力に比例する強さの地衡風が発生する. 気圧傾度力は気圧の高い方から低い方へ向かって等圧線に直角に働き, コリオリの力は北半球では風の進行方向右向きに (南半球では左向きに) 働く. それゆえ地衡風は等圧線に沿って北半球では気圧の高い方を右手に (南半球では左手に) 見る向きに吹く. 中緯度上空では, この地衡風と実際の風の間で 10 パーセント程度の相違がある. この相違分が大気に収束・発散を生み, さらに鉛直流を生み, この鉛直流により雲ができ雨が降る.

3.2 温度風

温度風とは大気内の温度差が原因として生じる, 高度の異なる 2 点での地衡風の差分ベクトルのことである [9]. あくまで差のことであり, 実際にそのような風が吹いているわけではない. 気圧が一樣であるが気温は一樣ではない平面を考えたとき, 気圧が一樣であるから気圧傾度力は働かず, この平面上では風は吹かない. 一方で静水圧平衡の関係式から気温が低いところほど高度が上がった時の気圧の下がり方は大きい. そのため, その平面上空では気温の高いところの上空が高気圧, 気温が低いところの上空が低気圧となり, 地衡風の関係より北半球では気温の高いところを右手 (南半球では左手) に見るように風が吹く.

3.3 温度風方程式

温度風方程式は図 1 で示すように異なる二点間の気圧 p_1, p_2 ($p_1 < p_2$, すなわち高さ高度では, p_1 が上で p_2 が下) を, その二高度間の平均気温 \bar{T} を用いて, 地衡風の差分ベクトル V_T で表される [10].

$$\{bfseries V_T = \frac{R}{f} \ln \left[\frac{p_1}{p_2} \right] \mathbf{k} \times \nabla_p \bar{T} \quad (1)$$

R は乾燥空気の比気体定数, f はコリオリパラメータを表す. これを x 成分の地衡風の速度ベクトル u , y 成分の地衡風の速度ベクトル v に分解して下の高度の速度ベクトルを移行すると, 高度に対する地衡風の差分式をえる. ここで, g は地衡風をさす.

$$u_t = u_{g_{i-1}} - u_{g_i} = -\frac{R}{f} \ln \left(\frac{p_i}{p_{i-1}} \right) \frac{\partial \bar{T}}{\partial y} \quad (2)$$

$$v_t = v_{g_{i-1}} - v_{g_i} = \frac{R}{f} \ln \left(\frac{p_i}{p_{i-1}} \right) \frac{\partial \bar{T}}{\partial x} \quad (3)$$

対流圏下層の観測データの信頼性は高いと仮定する. また一般に温度の観測データは, どの高度においても信頼性が高い. この温度風の関係をもとに, 対流圏の下層状態から上層の風を予測する予測式を次のように構築する.

$$u_{p_{i-1}} = u_{p_i} - \frac{R}{f} \ln \left(\frac{p_i}{p_{i-1}} \right) \frac{\partial \bar{T}}{\partial y} + A \quad (4)$$

$$v_{p_{i-1}} = v_{p_i} + \frac{R}{f} \ln \left(\frac{p_i}{p_{i-1}} \right) \frac{\partial \bar{T}}{\partial x} + B \quad (5)$$

A は水平風, 東西成分中の非地衡風成分, B は南北成分中の非地衡風成分を表す. 地表面付近の境界層より上空, 自由大気中では地衡風近似がよく成り立つ. そこで本研究では, 875 hPa を最下層とし, 上空の対流圏内 (250 hPa 程度) までの予測とし, $A=0, B=0$ としてこの予測式を使用した.

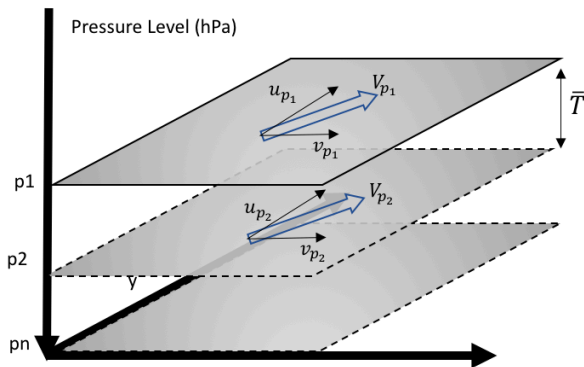


図 1 各気圧の地衡風の概略図

3.4 モデル

本研究では, 畳み込みニューラルネットを使用し, 湿度など観測データの特徴量から, 上空部の風を予測した. (1) で示した温度風方程式を畳み込みニューラルネットの損失関数に統合し, 教師あり学習を行う.

3.5 損失関数

損失関数には, x 成分と y 成分で表された温度風の更新式を用いる. X を学習データの特徴量, t を時間, f を畳

み込みニューラルネットからの出力値として損失関数を以下の様に定義する.

$$L(X_t, f(X_t)) = \sum_{i=1}^n |V_T - f(X_t)| \quad (6)$$

$$L_u(X_t, f(X_t)) = \sum_{i=1}^n |f_{i-1}(X_t) - \frac{R}{f} \ln \left(\frac{p_i}{p_{i-1}} \right) \frac{\partial T}{\partial y} - f_i(X_t)|$$

$$L_v(X_t, f(X_t)) = \sum_{i=1}^n |f_{i-1}(X_t) + \frac{R}{f} \ln \left(\frac{p_i}{p_{i-1}} \right) \frac{\partial T}{\partial x} - f_i(X_t)|$$

ここで, i は高度を表す. モデルからの出力値は全ての高度での x, y 成分をもつ風速度ベクトルを表す. i は高度を示しており, 図 2 で表す様に, 各高度での出力値と温度風の方程式から得られた差分により, 上空の値を推定する.

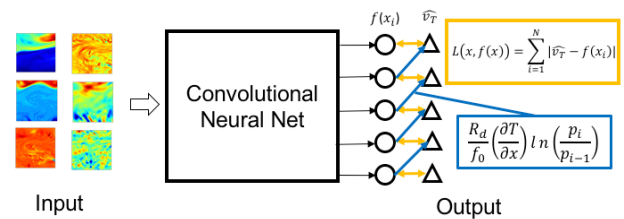


図 2 全体のアーキテクチャ

4. 実験

物理モデルと機械学習モデルを統合することによる有用性を検討するため, 実験を行なった. 以下に, 使用したデータセット, 比較手法, 結果, 考察についてまとめる.

4.1 データセット

データセットには, ERA-Interim を使用した. ERA-Interim は全世界の気象に関する様々な変数からなるデータセットであり, データ同化を用いて取得されている.

ERA-Interim を生産するために使用されたデータ同化システムは 2006 年の IFS (Cy31r2) に基づいている [11]. このシステムは, 12 時間の分析ウィンドウを備えた 4 次元変分分析 (4D-Var) を含む. データセットの空間分解能は表面から 0.1hPa までの 60 垂直レベルで約 80km (T255 スペクトル) である. ERA-Interim は完全な観測データではないが, 本研究は物理モデルと機械学習モデルの統合に関する初期検討であるため, 余分なノイズ等を排除するためにこのデータセットを使用した.

使用した特徴量を表 1 に示す. 一カ月の月平均データを使用しており, 学習データとして 1979 年 1 月から 2010 年 12 月までを使用し, テストデータとして 2011 年 1 月から 2016 年 12 月までを使用した. また, 本実験では太平洋上の北緯 15 度~北緯 75 度, 西経 150 度~東経 150 度の範囲のみを使用した. 解像度は経度・緯度共に 0.75 度

表 1 用いた特徴量および単位

	単位
速度ポテンシャル	s^{-1}
渦位	$m^2 \cdot s^{-1} \cdot K \cdot kg^{-1}$
湿度	kg/kg
風の鉛直成分	$Pa \cdot s^{-1}$
渦度	s^{-1}

表 2 各モデルと ERA-Interim の RMSE

	u	v
Physics-based	2.639451	3.289510
ERA-Interim based	2.321844	3.131168

であり、空間サンプル数は 60×60 点とした。高さ方向は 875hPa~250hPa までの 20 点を使用した。

4.2 比較手法

物理モデルと統合したニューラルネットの比較手法として、ERA-Interim の風の速度を教師信号として訓練したニューラルネットを使用する。また、ハイパーパラメータやネットワーク構造は全て同じものを用いた。ERA-Interim の実数値、物理モデルと統合したニューラルネットの予測値、ERA-Interim を教師信号とし訓練したモデルの予測値について、Root Mean Squared Error により比較する。

4.3 実験結果

ERA-Interim の実数値と物理モデルと統合したニューラルネットの予測値、ERA-Interim の実数値と ERA-Interim を教師信号とし訓練したモデルの予測値の Root Mean Squared Error の値を表 2 に示す。物理モデルを学習したモデルの方が RMSE の値が大きくなっていることがわかる。

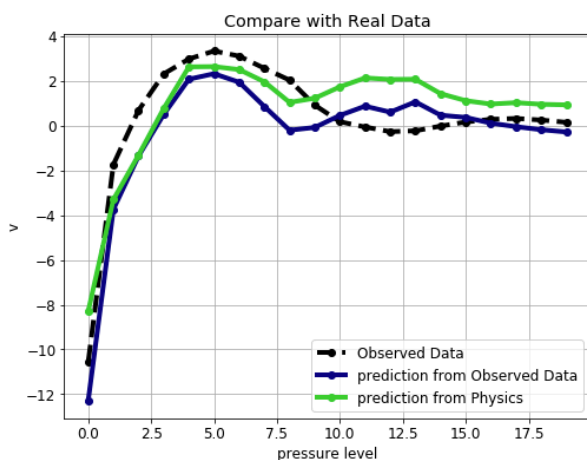


図 3 ある時刻 t における各モデルの出力値と ERA-Interim の比較

ある時刻 t における x 成分の各モデルの出力値および ERA-Interim の値の比較を図 3 に示す。双方データの特

徴を捉えた予測は行えているが、ERA-Interim から学習したモデルの方が高い精度で予測できていることがわかる。また、 x 成分の風ベクトルは、高度 400hPa 以上の帯域で、ジェット気流と呼ばれる気流が生じているため、一般的に大きな値になっている。この特徴が現れている時刻の予測は比較的うまく行われているが、この特徴が現れていない時刻についてはほぼ学習できていないことがわかった。

4.4 考察

物理モデルを学習したモデルでも風の特徴を捉えた学習ができていることがわかる。しかし、物理モデルを学習したモデルの方が RMSE の値が大きくなった。今回予測した風の速度は、対流圏上層の風を予測しているため地衡風に近似しているが、それでも非地衡風成分を無視できるほどではない。この非地衡風成分が精度の差として現れたのではないかと考えられる。今後はこの誤差分も機械学習で推定することも重要になってくると考えられる。

また、今回は、教師信号に観測された風ベクトルの値を使用していない。この手法でも、ある程度の予測を行えるが、教師信号を明示的に与えた場合より精度よく予測することは難しい。精度よく予測を行うには、物理モデルのみから学習させるのではなく、観測データと物理モデル双方を教師信号として損失関数に統合する必要があると考えられる。

5. まとめ・今後の展望

本研究では、自然科学分野で使用される物理モデルを機械学習モデルに統合することにより、機械学習の問題点を解決する手法に着目した。大気内の温度差を原因として生じる地衡風の速度の高度差である温度風を予測することを初期検討の例として使用した。東西成分と南北成分からなり、差分ベクトルとして表される温度風方程式を制約としてニューラルネットワークの損失関数に加え学習を行う手法を実問題に近い形で適用し、その有用性を検討した。結果から、一定の精度で予測を行うモデルを構築することができた。

今後は、非地衡風成分などの誤差分も機械学習で推定できるように拡張していく。また、物理モデルと観測データ双方を教師信号として使用し、既知の法則性を無視した出力を得るといった問題点を解決できる手法の開発を行う。

その後は、物理モデルにより、ニューラルネットワークの高次元のパラメータの範囲を限定させ、物理法則に沿った学習のみを行える解釈性・説明性の高い手法の検討を行う。

そして、時系列データの予測や、Navier-Stokes 方程式と機械学習を統合させるなど、より貢献度の高い問題を扱うことを目的としていく。

参考文献

- [1] Lazer, David and Kennedy, Ryan and King, Gary and Vespignani, Alessandro: *The Parable of Google Flu: Traps in Big Data Analysis*, pp.1203-1205, *Science* (2014).
- [2] Marco Tlio Ribeiro, Sameer Singh and Carlos Guestrin: "Why Should I Trust You?": *Explaining the Predictions of Any Classifier*, pp.1135-1144, *KDD* (2016).
- [3] Anuj Karpatne, Imme Ebert Uphoff, Sai Ravela, Hassan Ali Babaie and Vipin Kumar: *Machine Learning for the Geosciences: Challenges and Opportunities*, *arXiv* (2017).
- [4] Grgoire Montavon, Wojciech Samek and Klaus Robert Miller: *Methods for Interpreting and Understanding Deep Neural Networks*, *arXiv* (2017).
- [5] A. Karpatne, G. Atluri, J. H. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, S. Shekhar, N. Samatova and V. Kumar : *Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data*, pp.2318-2331, *IEEE TKDE* (2017).
- [6] Breiman Leo: *Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author)*, pp.199-231, *Statist. Sci.* (2001).
- [7] Russell Stewart and Stefano Ermon: *Label-Free Supervision of Neural Networks with Physics and Domain Knowledge*, pp.2576-2582, *AAAI* (2017).
- [8] Jonathan Tompson, Kristofer Schlachter, Pablo Sprechmann and Ken Perlin: *Accelerating Eulerian Fluid Simulation With Convolutional Networks*, *ICML* (2016).
- [9] 小倉 義光: *総観気象学入門*, 東京大学出版会 (2000).
- [10] John Marshall, Alan Plumb and Lodovica Ilari: *Thermal wind*, MIT PAOC. (2003).
- [11] Paul Berrisford, Dick Dee, Paul Poli, Shinya Kobayashi, Sakari Uppala and Adrian Simmons: *The ERA-Interim archive*, *ECMWF*. (2011).