機械学習による天気予報の当たり外れ予測の可能性 How can we predict the accuracy of weather forecast with machine learning?

船木将秀* Masahide Funaki

*1 新潟大学創生学部 College of Creative Studies, Niigata University

A better weather forecast (WF) is of public interest. Over the past decades, the precision of WFs has been significantly improved by developments of more advanced forecasting techniques with state-of-the-art numerical models and machine learning. However, validation results of such precision hardly go public, and thus individual WF user tends to evaluate the accuracy of every WF through his 'rules of thumb'. Predictive values of a forecast inevitably vary with changes in temporal-spatial dynamics of the atmosphere which is simulated by a set of numerical models that have both strength and weakness in their performances. Largely using web-scraping, we examine the causality of simulated atmospheric dynamics (inputs) and predictive values (outputs) through deep learning, in order to evaluate the accuracy of WFs in a daily basis. It is anticipated that our evaluation results will allow users to choose better forecasting systems.

1. はじめに

天気予報は、複雑な気象系の解析・予測の結果であり、それ は社会活動に大きな影響を及ぼすものである.したがって、予 報精度向上のために、気象庁や民間の予報業務許可事業者を 中心に、日々様々な取り組みがなされている.例えば、観測デ ータの量的・質的向上や解析手法・計算能力の高度化に伴っ て、気象庁の数値予報は、1959年のコンピュータ導入以降、 年々予報精度は向上している[気象庁 2018].

また,近年では,機械学習による気象予測をする研究[気象 研究所予報研究部 2000][吉兼 2018]もなされている.実際の 予報業務においても,機械学習の利用は進められている.気象 庁では,予報ガイダンスの導出に機械学習を用いて気温予報 精度に努めており[高田 2016],民間の予報業務許可事業者 でも,機械学習の導入により,予報精度が大幅に向上した例が 報告されている[ウェザーニュース 2018].

天気予報の精度全体は年々向上している一方で,予報精度 が定量的に分析・公表されることは少ない.利用者は利用する 予報が実際に当たるかどうかを感覚的にしか理解できないのが 現状である.

本研究では、現時点での天気予報の当たり外れを、天気予報の得手不得手という観点から機械学習で予測するシステムの 構築に当たり、その可能性について議論する.

2. 天気予報の適中精度

2.1 予報における誤差

予報精度が年々向上しているとはいえ、大気の状態を完全 には表現できない以上、予報には常に誤差が含まれる.この誤 差は、予報モデルと現実との差異によって生じるものである.気 象庁の天気予報を例に挙げると、数値予報で大気の状態をシミ ュレートするときと、天気予報で数値予報を解釈するときに、そ れぞれ誤差が発生している.

(1) 数值予報

数値予報では、大気の支配方程式を積分することで未来の 大気の状態を予測する.この際に用いる計算モデルは、気象機 関や予報内容によって解像度や計算処理内容が異なっている. モデルの違いは予報内容に影響を与え、例えば熱帯低気圧の 進路予報では、それぞれが異なった予報を出している [Yamaguchi et al. 2017].

数値予報の誤差が出る要因は3つに集約される.1点目は計 算モデルの入力値に観測誤差が含まれることである.2点目は 数値予報の初期値は、客観解析などの近似計算による誤差が 含まれるということである.現実的にこれら2点の根本的な解決 は不可能であり、数値予報モデルが完全に大気の状態をシミュ レートできないということである.3点目は、初期値の誤差が時間 の経過とともに拡大することである.一般に、大気の支配方程式 は非線形であり、数値予報の結果はカオス理論的な振る舞いを 示すことが知られている[Lorenz 1963].

それぞれのモデルの違いは、いわばモデルごとの得手不得 手であり、モデルの癖とでも言うべきものである。2004 年~2014 年における全球での熱帯低気圧の3日後の進路予測と実際の 中心位置との誤差について、各気象機関で大きな差が出てい る(図1).



図1熱帯低気圧3日後予測進路と実際の中心位置

連絡先:船木将秀,新潟大学創生学部,新潟市西区五十嵐2 の町 8050 番地, x18c064j@mail.cc.niigata-u.ac.jp

との誤差(出典: [Yamaguchi et al. 2017])

(2) 予報ガイダンス

予報ガイダンスは,数値予報の結果(大気の状態)を翻訳・修 正して得られる予測資料である.数値予報の翻訳・修正の過程 では,数値予報での予測値を補正し誤差を軽減したり,天気や 発雷確率などの数値予報で出力しない情報を作成したりする.

予報ガイダンスでは、統計的な手法を用いて、数値予報の誤 差や数値予報にない予報データを導出する.気象庁では、線 形重回帰のロジスティック回帰ほか、ニューラルネットやカルマ ンフィルターなどの機械学習の導入も行われている.

(3) 天気予報

天気予報は、数値予報や予報ガイダンスなどの予測資料と、 気象衛星や AMeDAS などの観測資料をもとに、予報官や気象 予報士が作成している.この過程で、それぞれの担当者の経験 や知識といったものが加味されている.また、ウェザーリポート [ウェザーニュース 2018]のように、追加で独自のデータを加味 することもある.発表予報を作成する過程では、担当者の経験 や知識に予報精度が左右される側面がある.

また,それだけではなく,予報範囲の問題もある.気象庁は 2019年1月現在142の予報区で予報を行っている.一方で, 日本気象協会では,市区町村別の予報を発表している.この二 者を比較したとき,前者のほうが予報する区域が広く,予報区内 で天気にばらつきがあるような,不安定な大気の状態であれば, 必然的に前者のほうが的中率は低くなる.

2.2 予報精度の現況

予報精度について、気象庁では月ごとに前月の各予報区・ 地方ごとに、天気予報の予報精度の検証結果を公表している [気象庁 2019]. この検証では、気象庁 5 時・17 時発表の天気 予報の、降水の有無、最高・最低気温の予報データと観測デー タとを比較し、1 ヶ月単位で集計している. この検証データによ れば、長期的な精度の向上はあるものの、地域や季節によって 予報精度にばらつきがある(図 2).

前述のように、冬(12 月~2 月)の北海道の降水の有無の予 報精度は全国平均と比較して低い.また、全国的に夏(7 月~8 月)の降水の有無の予報精度も低くなる傾向がある.このような 天気予報には季節性や地域性を帯びた予報精度のばらつきは、 前述の 2 つの誤差に起因するものである.つまり、予報の難易 度は、大気の状態のシミュレーションの難易度に依存し、それは、 予報精度の地域性や季節性と言ったパターンで表面化してい る.

ところが、この季節性や地域性は、あくまで平均的な傾向であり、気象状況によって、予報の難易度が変化することもある. 2018年と平年値(1992年~2017年の平均)を比較すると、2018年の降水の有無の適中率は、差分が0~3%程度が続くなど、概ね平年値と同様に推移しているが、多くの地域が高気圧に覆われ大気の状態が安定していた期間が長かった3月、7月の適中率は平年値より9%程度高かった(図3).



図 2 気象庁翌日予報(17時発表)の降水の有 無の適中率・平年値(1992~2017年の平均値)



図 3 気象庁翌日予報(17時発表)の降水の有無 の適中率の平年値(1992年~2017年)と2018年 との比較



図4 各観測地点での適中率(イメージ)

3. 機械学習による天気予報の当たり外れ予測

天気予報の当たり外れは、従来、感覚的な経験則や気象庁 の予報検証のような統計分析によるものであった.しかし、天気 予報の予報精度には規則性があるとはいえ、その要因は複雑 であり, 現時点での予報が, 当たるか外れるかは予測することが 困難であった. しかし, 深層学習を利用することで, 人間に認識 できないパターンを発見することが期待される.

3.1 機械学習での当たり外れ予測の概要

本システムでは,各予報時点での気象状態(数値予報の初 期値)と予報地点・時刻ごとの天気予報を入力とし,その予報の 当たり外れを出力とすることを想定している.過去の予報の当た り外れに関しては,予報の内容と当該時点の AMeDAS などの 気象観測網でのデータとの比較を行うが,twitter での天気に関 するツイートを収集し,地点情報と紐づけて天気の実況データと して扱うことを構想している.これは,AMeDAS の間隔が降水で 約17km間隔,気温や風で約21km間隔であることや,全天の8 割が雲で覆われていても晴れである場合のように,観測値と体 感の違いがあると考えられるため,より一般利用者の感覚に合 った予報適中予測をするためである.

これを,予報者(気象機関・予報許可事業者など)ごとに学習 することで,大気の状態や予報内容と,予報適中率との規則性 をニューラルネットワークが獲得する.



図 5 天気予報の当たり外れ予測システムの構想 図

3.2 機械学習での当たり外れ予測の課題点

機械学習の実装において、重要となってくるのは、学習する データの量・質を必要十分に設定することである.気象のような 複雑でカオス性を持つものは特に、適切な種類で十分な量の データを吟味して、収集することが必要である.しかし、過去の 予報データは、それ自体には何ら価値はなく、振り返って収集 することが困難である.したがって、天気予報の当たり外れを予 測するにあたっては、十分な量の予報データを集められるかが 課題となる.

3.3 当たり外れ予測への期待

2019年1月現在、日本には、気象・波浪の予報業務許可事 業者が74者存在し、そのすべてが天気予報の業務を行うことが できる.また今後、解析精度や計算能力の向上、機械学習の利 用等で、更に天気予報の質・量ともに増大することが見込まれる. こうした中で、各種天気予報の当たり外れを予測ができれば、 利用者が目的に応じて、最も可能性の高い予報を利用すること ができることが期待される.これは、今まで感覚的に行われてき たことであるが、定量化することで、広く一般に、より正確な予報 を得られる様になると考えられる.また、機械学習で一定以上の 品質が担保できれば、機械学習による逐次的な予測をすること で、統計分析による予報精度の定量的比較より現時点での予 報の当たり外れに即した評価ができるだろうと期待される.

謝辞

本研究に取り組むに際しまして、予報データの分析や研究資料の検索に協力してくださった、新潟大学創生学部・峰川拓人さん、齋藤季さんによって、本稿は完成することができました。ありがとうございました。

参考文献

- [Lin et al. 2018] Lin, S.-Y., Chiang, C.-C., Lic, J.-B., Hunga, Z.-S., Chao, K.-M.: Dynamic fine-tuning stacked autoencoder neural network for weather forecast, Future Generation Computer Systems Volume 89, 446-454, The International Journal of eScience, 2018.
- [Lorenz 1963] Lorenz, E.N.: Deterministic Nonperiodic Flow, Journal of the Atmospheric Sciences Vol.20, 130-141, AMS, 1963.
- [Yamaguchi et al. 2017] Yamaguchi, M., Ishida, J., Sato, H., Nakagawa, M.: WGNE Intercomparison of Tropical Cyclone Forecasts by Operational NWP Models: A Quarter Century and Beyond, Bulletin of the American Meteorological Society 98, AMS, 2017.
- [気象研究所予報研究部 2000] 気象研究所予報研究部:帰 納的機械学習手法の天気予報への応用,気象研究所技 術報告 第39号,2000.
- [気象庁 2018] 気象庁: 数値予報の精度向上, 2018, https://www.jma.go.jp/jma/kishou/know/whitep/1-3-9.html.
- [気象庁 2019] 気象庁: 天気予報の検証結果(平成 30 年の年 集計), 2019, https://www.data.jma.go.jp/fcd/yoho/data/kensho/HPdata201 8/zc_2018.html.
- [国次 1997] 国次雅司: カルマンフィルターを用いた天気予報 ガイダンス, 天気 44(6), 413-417, 日本気象学会, 1997.
- [松村 2006] 松村崇行: 天気予報ができるまで, 細氷 52 号, 日本気象学会北海道支部, 2006.
- [三隅 2014] 三隅良平:気象災害を科学する、ベレ出版、2014.
- [饒村 2002] 饒村曜: 気象災害の予測と対策, オーム社, 2002.
- [小倉 2015] 小倉義光:日本の天気 その多様性とメカニズム, 東京大学出版会, 2015.
- [小倉 2016] 小倉義光: 一般気象学, 東京大学出版会, 2016.
- [高田 2016] 高田伸一: 気象庁における機械学習の利用, 先端 IT 活用推進コンソーシアム, 2016, http://aitc.jp/events/20160916-Seika/20160916 特別講演 気象庁における機械学習の利用.pdf.
- [ウェザーニュース 2018] ウェザーニュース:目指せ精度 99.9%, ウェザーニュースの予報がわずか 1 年で大幅に改善した前 代 未 聞 の プ ロ ジ ェ ク ト , 2018 , https://weathernews.jp/s/topics/201703/220235/.
- [山下 2018] 山下隆義: イラストで学ぶ ディープラーニング, 講談社サイエンティフィク, 2018.
- [吉兼 2018] 吉兼隆生,芳村圭:機械学習による局地気象予測の試み,人工知能学会全国大会論文集 第32回全国大会,人工知能学会,2018,
 <u>https://www.jstage.jst.go.jp/article/pjsai/JSAI2018/0/JSAI20</u>18_4Pin134/_pdf/-char/ja.