

NGBOOST を用いた日射量予測の基礎検討

Basic study of predicting solar radiation using NGBOOST

池田 健一郎[†] 浦野 昌一[†]
Kenichiro Ikeda Shoichi Urano

1. はじめに

近年、人為起源による温室効果ガスが年々増加し、地球温暖化が深刻化している背景において、世界各国では 2050 年に向けて温室効果ガスの排出を全体としてゼロにするカーボンニュートラルが目標として掲げられている。温室効果ガスが排出される原因の一つとして、化石燃料の利用が挙げられている。日本の電力供給における電源構成では化石燃料による火力発電が大きく占めていることから、温室効果ガスを排出しない再生可能エネルギーの活用が注目されている。特に、太陽光発電は設置に制限がなく、導入がしやすい電源システムであるため導入が促進され、年々増加傾向である。太陽光発電を運用していく際には火力発電や水力発電等の他の電源システムと調和して電力供給する必要があり、発電計画を立てる場合に需要量や発電量の予測が必要とされる。予測が大きく外れた場合は、電源に余剰もしくは不足が発生し需給のバランスが崩れ停電に繋がる可能性が考えられる。そのため、電力の安定的な供給には、高精度な予測が必要になる。

太陽光発電量予測では、高精度の予測値算出の実現のために機械学習を用いた研究が行われている。太陽光発電は天候により発電量が左右され不確実性が伴うため、予測の高精度に加え不確実性を考慮する手法が望まれている。また、精度向上のため、気象状態を考慮した研究が行われており、日射量予測では先行研究^[1]において、複雑な気象変化を定量的に表現することで精度向上を実現している。しかし、機械学習を用いた気象分類に基づく日射量予測には検討には研究の余地がある。

筆者らは、これまで気象データにクラスタリング手法を適用することで気象状況を分類し、XGBOOST による日射量予測を行ってきた^[2]。研究の結果、気象変化の細かい要素を分類する事が出来たが、日射量予測の有効性は確認出来なかった。問題点として、クラスタリングを行う際の入力データとクラスタ数の選定が挙げられた。

そこで、本稿では不確実性を考慮する事が可能である NGBOOST を日射量予測に適用しその有効性を検証する事に加え、クラスタリング手法による気象分類のさらなる検討を行うことで日射量予測に適したモデル構築を目指す。

2. 日射量予測

本稿では、予測対象として、1 時間ごとの積算日射量を予測する。日射量は、時々刻々と変化する気象変化によって短い時間間隔で日射量に変化する特徴を有する。そのため日射量予測を短い時間間隔で行う必要がある。一般的に、太陽光発電量を予測する場合、直接的に太陽光発電量を予測する方法もあるが、日射量は太陽光発電量と強い相関があるため日射量を予測することで間接的に太陽光発電量を算出する方法もある。

[†] 明治大学先端数理科学研究科 Meiji University
Advanced Mathematical Sciences

3. 予測手法

筆者らは先行研究^[2]において予測手法として XGBOOST 用いた日射量予測を行ってきた。本稿では、予測の不確実性を考慮する事ができる NGBOOST を適用し先行研究との比較を行う。また、気象分類のクラスタリングには k-means++を用いる。

3.1 k-means++

k-means++とは、クラスタリングするための機械学習手法であり、教師なし学習の一つである。従来の k-means 法の初期値依存問題に対して、k-means++ではクラスタの初期値同士を離れた点に設定する事で、できるだけ特異なクラスタを作らない事を実現している。

3.2 XGBOOST

XGBOOST は機械学習の教師あり学習に分類され、回帰、分類問題を解く事ができる手法である。アンサンブル学習のブースティングの一つであり、勾配ブースティング手法の派生型である。過学習低減のためのアルゴリズムとして損失関数に正則化項が組み込まれている点特徴的である。

3.3 NGBOOST

NGBOOST は機械学習の教師あり学習に分類され、回帰、分類問題を解く事ができる手法である。勾配ブースティング手法の派生型のアルゴリズムであり、一般的な点推定では無く、確率分布を推定する事で予測点の不確実性を定量的に評価出来る点特徴的である。学習の最適化を行う際に、ユークリッド空間からリーマン多様体として一般化することで自然勾配法を用いる事が可能となり、確率分布予測を実現している。

4. シミュレーション

本稿では、XGBOOST と NGBOOST の比較検証した後、k-means++でのクラスタリング結果に基づいて NGBOOST

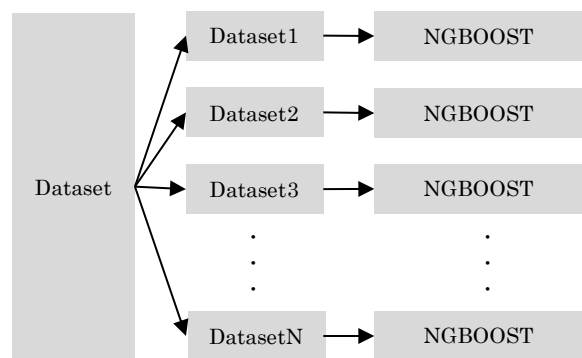


図 1 シミュレーションフロー

を用いた日射量予測を行う。シミュレーションの流れを図 1 に示す。

4.1 シミュレーション条件

4.1.1 入力変数

以下の 10 要素を入力変数とする。

気温、相対湿度、風速、雲量、降水量、気圧、蒸気圧、露点温度、時間要素、太陽高度
上記の観測地点は、日射量対象地点と最も近い東京地点の気象データを使用する。

4.1.2 モデル化期間

1 時間ごとの日射量予測を行う際に、今回は 2016 年 1 月～2018 年 12 月まで学習対象期間とし、2019 年 1 月～2019 年 12 月までを予測対象期間とする。

4.1.3 手法のパラメータ設定

分類には k-means++、予測には XGBOOST と NGBBOOST で予測する。その際の条件を以下に示す。

<k-means++>

クラスタ数：2~5

<XGBOOST>

学習率：0.3

決定木の数：1000

<NGBBOOST>

学習率：0.3

決定木の数：1000

予測分布：正規分布

4.2 シミュレーション結果

4.2.1 XGBOOST/NGBBOOST による精度比較

精度比較を行うため、2016 年~2018 年を学習、2019 年を予測した。平均誤差率、MAE、RMSE の結果を表 1 に示す。結果から、NGBBOOST は XGBOOST より全ての評価指標において精度が良い結果が得られた。また、NGBBOOST は XGBOOST と比べ、不確実性が考慮できる点から、日射量に対して有効的な手法であると確認できた。

4.2.2 クラスタ数の変化による精度比較

クラスタ数の変化による精度比較を行うため、クラスタ数を 2 から 5 まで変更し NGBBOOST により学習、予測を行った。k-means++ の入力変数には気温、相対湿度、風速、雲量、降水量、気圧、蒸気圧、露点温度を利用した。平均誤差率を基に気象庁ラベル毎に合わせた精度比較を表 2 に示す。今回、精度向上が確認された箇所はクラスタ数 3 の

表 1 手法の精度比較

	NGBBOOST	XGBOOST
平均誤差率[%]	45.485	47.020
MAE	0.1159	0.1194
RMSE	0.0867	0.0884

表 2 平均誤差率の精度比較

気象庁ラベル	クラスタ数=2	クラスタ数=3	クラスタ数=4	クラスタ数=5
快晴	21.733	23.292	21.998	22.716
晴	26.963	26.834	28.130	27.515
薄曇	27.131	27.256	28.998	27.835
曇	56.446	55.457	54.619	55.387
雨・雪など	76.135	103.489	101.387	98.368
全体	42.275	46.781	44.989	46.194

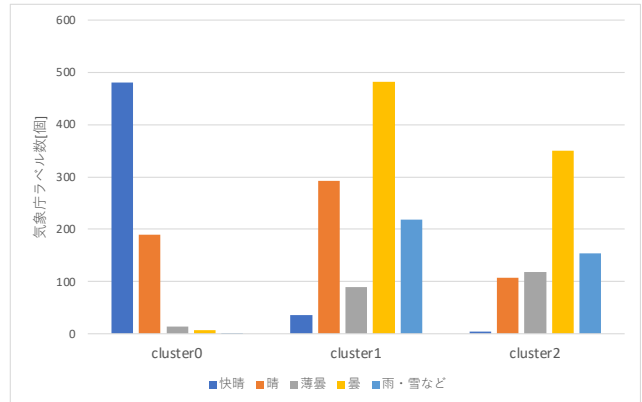


図 2 クラスタ毎に属する気象庁ラベル数

表 3 平均標準偏差の精度比較

	気象庁ラベル	クラスタ数=2	クラスタ数=3	クラスタ数=4	クラスタ数=5
快晴	0.0158	0.0324	0.0215	0.0220	0.0212
晴	0.0290	0.0475	0.0361	0.0375	0.0322
薄曇	0.0236	0.0616	0.0511	0.0514	0.0454
曇	0.0423	0.0572	0.0526	0.0513	0.0436
雨・雪など	0.0459	0.0332	0.0315	0.0288	0.0250
全体	0.0327	0.0467	0.0392	0.0388	0.0338

「曇」である。クラスタ数 3 に属する気象庁ラベルの個数を図 2 に示す。クラスタ数 3 における分類では、Cluster0 が晴要素の強いクラスターとして構成された。Cluster1 と Cluster2 では「曇」が二つに分割された事が分かる。Cluster1 では「晴」が多く含まれている事から、晴寄りの曇、一方 Cluster2 では雨寄りの曇であると考えられる。クラスタ数 4、5 では精度が悪くなっている事から、「曇」の特徴を捉えたモデル作成において、細かく分類するよりも、ある程度の柔軟性を持たせた分類が適切であるという事が分かった。平均標準偏差を基に気象庁ラベル毎に合わせた精度比較を表 3 に示す。クラスタ数を増やすごとに「雨・雪など」の平均標準偏差が減少している事が分かる。細かく分類することで予測の信頼性が増したが、誤って分類された部分で平均誤差率が大きく出てしまったのだと考えられる。

5. まとめ

本稿では、XGBOOST と NGBBOOST の精度比較に加え、k-means++ の結果を活用した NGBBOOST による日射量予測を行った。今回、NGBBOOST は日射量予測に対して有効的な手法である事が確認できた。また、k-means++ を用いたシミュレーションでは、クラスタ数の変化によって一部の要素では精度向上は確認できた。

今後の課題としては、精度向上が確認出来なかった要素の更なる検討として、気象データにソフトクラスタリングを適用した分析を行い、日射量予測の精度向上を目指す。また、他の確率分布を推定する手法との比較を行う事で、NGBBOOST の有用性を検証していく。

参考文献

- [1] 嶋田 尊衛,黒川 浩助,“天気予報と天気変化パターンを用いた日射量予測”, 電気学会論文誌 B, Vol.121, No11, pp1219-1225 (2007)
- [2] 池田 健一郎, 浦野 昌一,“クラスタリングによる気象分類に基づいた日射量予測の検討”, 2022 年度人工知能学会全国大会 (第 36 回), No. 2D6-GS-04 (2022)