

再生可能エネルギー発電実績データの
データインフラストラクチャー構築と発電量予測

小川滉介[†] 佐藤彰洋[‡]

横浜市立大学データサイエンス学部データサイエンス学科^{†‡}

はじめに

現在、温室効果ガスを排出しない再生可能エネルギーの普及が注目されている。特に、複数の電力を束ねるアグリゲーター[1]は、電力の需給バランスの安定化を図り、再生可能エネルギーを有効に活用するため、発電量の予測が重要な役割となっている。佐々木ら[2]は天候を数値化し、天気予報から太陽光発電出力推定手法を検証し、天候数値化係数と発電量には強い相関があり、天候数値化係数による太陽光発電出力推定が可能であることが確認されている。

本研究は、これらの背景と従来研究を参考に、過去の発電量実績と気象データを用いて、予測期間における天気から発電量を予測する手法を検討する。本研究では、再生可能エネルギーの流通により地域活性化を図る株式会社まち未来製作所[1]との共同研究として提供を受けた、発電所の発電実績データを元にデータインフラストラクチャーを構築し、各発電所における発電量予測システムの実装をめざした。

本研究の目的は、効率的な発電量の予測を可能にすることであり、これにより、電力取引における発電計画値の決定を支援できると考えられる。

2.1 データインフラストラクチャー

本研究では、データ収集から可視化、予測モデル構築までのプロセスモデルフレームワークに従い、データインフラストラクチャーの構築と発電所毎の発電量予測プログラムの開発を行った。太陽光発電所における発電量実績データと気象データの2種類のデータを用いて、データインフラストラクチャーを構築し、発電量予測プログラムの開発と評価を行う。発電量実績データでは、30分ごとの発電量がkWh単位で表示されている。気象データは、太陽光発電所付近の1時間ごとの気象データを気象庁公開データ[3]から取得した。気象データは1時間を30分に分割し同じ状態が継続するとした。

学習データとして2022年11月のデータを使用して発電量予測プログラムを開発し、評価データとして2023年11月のデータを使用して発電量予測プログラムの評価を行う。ただし、2023年11月8日と9日と17日は発電所の点検日であったため評価の対象外としている。

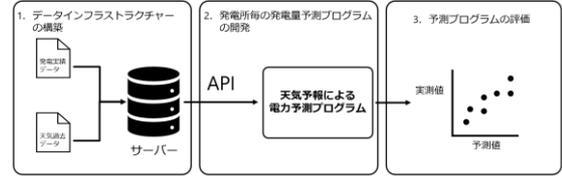


図1: 分析システムの処理の流れ

変数	説明
n	予測期間におけるデータ数
t	予測期間より1年前の時間帯 (1~48)
w	予測期間より1年前の天気
$\widehat{D}_{t,w}$	予測期間より1年前の発電量における集合
T	予測期間の時間帯(1~48)
W	予測期間における天気
$D_{T,W}$	予測期間における実績発電量の集合
$\widehat{G}_{T,W}$	1年前のデータによる特定の時間帯と天気における発電量の集合 (昇順に並び替え)
n_G	$\widehat{G}_{T,W}$ のデータ数
$q2_{T,W}$	$\widehat{G}_{T,W}$ における発電量の中央値 (予測値)
MAE	実績発電量と中央値による平均絶対誤差
RMSPE	実績発電量と中央値による平均平方二乗誤差率

表1: 発電量予測における変数一覧

2.1 データインフラストラクチャー

図1に分析システムの処理の流れを示す。発電量実績データと気象データは、自動整形プログラムとスクレイピングプログラムを作成し、データベースとして自動保管するプログラムを開発した。また、各データを自由に取り出せるようにMESHSTATS[4]にWeb APIを作成し、電力量予測プログラムを開発する中で必要なデータの自動抽出を行い、効率的な発電量の予測を可能にした。

2.2 発電量予測プログラムの開発

$$\widehat{G}_{T,W} = \{\widehat{D}_{t,w} | t = T, w = W\}_{sorted} \quad (1)$$

$$q2_{T,W} = \widehat{G}_{T,W}^{(n_G+1)} \quad (2)$$

$$q1_{T,W} = \widehat{G}_{T,W}^{(3(n_G+1)/4)} - \widehat{G}_{T,W}^{(n_G+1)} \quad (3)$$

表1に発電量予測の変数を示す。(1)式~(3)式で示すように、まず予測期間以前における発電量実績と気象のデータを用いて、各時間と天気における発電量の四分位数とを算出する。そして、予測期間における各時間の天気と一致する発電量の中央値を予測値、四分位範囲を予測範囲として扱う。

発電量予測の評価として、2つの方法を用いる。

$$MAE = \frac{1}{n} \sum_{i=1}^n |q^2 \widehat{r_{i,W_i}} - D_{T_i,W_i}| \quad (4)$$

$$RMSPE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{(D_{T_i,W_i+1}) - (q^2 \widehat{r_{i,W_i+1}})}{(D_{T_i,W_i+1})} \right)^2} \quad (5)$$

1つ目は、予測期間における各日の実績値と予測値の誤差を使い、(4)式で定義される平均絶対誤差(Mean Absolute Error; MAE)と、(5)式で定義される平均二乗パーセント誤差(Root Mean Squared Percentage Error; RMSPE)を評価指標とし、予測における精度を評価する。2つ目は、予測期間における各時間帯のMAEと予測範囲(四分位範囲)の相関係数を算出し、予測範囲による予測精度を評価する。

3. 予測結果

2023年11月における発電量予測の評価結果を図2,3に示す。図2からは、日によって予測精度がばらついており、晴れの日には評価指標が低く、曇や雨の日には評価指標が高くなった。しかし、RMSPEは実績発電量が小さく、予測から外れていると高く計算される。そのため、日の出と日の入りの時間帯で予測から大きく外れている日であると、RMSPEが高くなっていると考えられる。図3からは、MAEと四分位範囲の相関関係は正の相関を有していた。すなわち、予測値の四分位範囲は予測精度を示しているといえる。また、回帰直線から外れているMAEが極端に大きな値の点は観測事象が少ない曇や雨の場合に該当している。

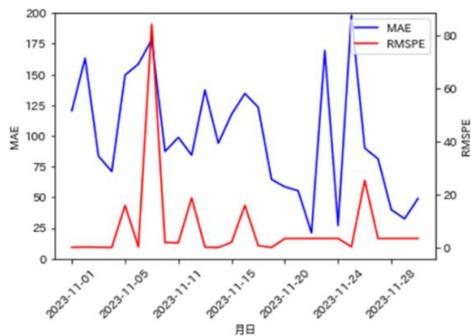


図2：各日におけるMAEとRMSPEの推移

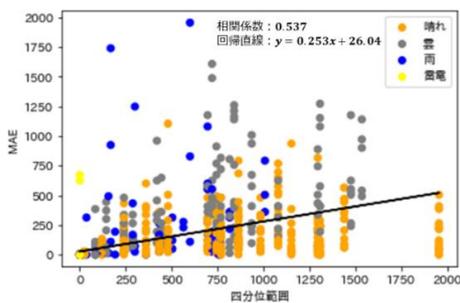


図3：MAEと四分位範囲の相関関係

これらの結果から、予測期間における各時間帯の天気と一致する発電量のデータがあるほど、予測精度が向上すると考えられる。そのため、晴れなどの良く起きる天気での予測には提案モデルの予測精度が良かった。しかし、曇や雨などのあまり起きない天気での予測は、該当するデータの数が少なく、予測精度は悪かった。曇や雨などの発電量実績データを複数発電所、および、長期間にわたり集めることで、精度向上を行う必要がある。

4. まとめ

本研究では、データインフラストラクチャーを構築し、過去の発電量と天気を用いて、予測期間における各時間帯の天気からの発電量を予測するプログラムを開発し、評価を行った。データインフラストラクチャーでは、発電量実績データと天気データを自動成形し、Web APIを新しくMESHSTATS上に開発したことで、効率的な発電量の予測を可能にした。

発電量予測プログラムでは、晴れなどの高頻度条件での予測精度は高く、曇や雨などの低頻度条件での予測精度は悪かった。

今後の課題として、将来の天気予報を本システムに組み込むことで、発電量予測を行うプログラムの作成がある。また、空間分解が高いメッシュ型気象データを利用することで予測精度を高める改善が必要である。

謝辞

本研究は横浜市立大学と株式会社まち未来製作所の共同研究「e.CYCLE[5] × データサイエンス」で作るサステナブルな未来都市(2022年11~2024年3月)として実施された内容を含みます。本研究にご協力いただいた株式会社まち未来製作所の皆様へ深くお礼申し上げます。

参考文献

- [1] 株式会社まち未来製作所：事業内容, <https://machimirai.co.jp/service.html>, (最終アクセス年月日: 2023年12月31日)
- [2] 佐々木三郎, 福永青空, 太田豊, 天候数値化簡易手法による太陽光発電の地域別および全国大の発電量の解析, エネルギー・資源学会論文誌, Vol. 38, No. 5 (2017), pp. 27-35. https://www.jstage.jst.go.jp/article/jjsr/38/5/38_27/_pdf/-char/ja
- [3] 気象庁：過去の気象データ <https://www.data.jma.go.jp/gmd/risk/obsdl/index.php>, (最終アクセス年月日: 2023年12月31日)
- [4] MESHSTATS, <https://www.meshstats.xyz/meshstats/>, (最終アクセス年月日: 2023年1月6日)
- [5] 株式会社まち未来製作所：e. CYCLE, <https://ecycle.net/>, (最終アクセス年月日: 2024年1月10日)