

数値気象モデル WRF とスパース回帰に基づくエリア太陽光発電量予測技術

Day-ahead Forecasting of Photovoltaic Power Generation by WRF and Sparse Regression

進 博正*・志賀慶明*・柿元 満*

1. はじめに

我が国における太陽光発電システムの導入容量は、固定価格買取制度（FIT：Feed-in Tariff）の導入や設備コストの低下を背景として年々増加している。たとえば、IEA（International Energy Agency）は太陽光発電や風力発電を変動する再生可能エネルギー源（VRE：Variable Renewable Energy）と定義して、VRE 導入量を4段階（第1段階～第4段階）に分けた上で、各段階で考慮すべき事柄を整理している⁽¹⁾。IEAによると我が国は、VRE 導入量（主に太陽光発電）の第2段階（九州エリアは第3段階）にあり、電力システム運用時にVREの影響を無視できずVREの出力予測が必要な段階に達している⁽¹⁾。太陽光発電量の予測は、電気事業者の計画業務用途で既に実用化しているが、最新技術を取り入れた精度向上も続いている⁽²⁾。Carloらは、系統運用者の立場から太陽光発電量予測の経済的価値を評価しており、精度向上が経済的メリットに直結すると報告している⁽³⁾。またHongら、加藤は、米国及び日本における太陽光発電量予測を含むコンテストの開催結果を報告している^(4,5)。このように太陽光発電量予測の精度向上は、経済的メリットに直結するデータ科学的に意義ある課題であり、オープンイノベー

ションを活用したコンテストに適した課題である。我が国でも東京電力ホールディングス殿と北海道電力殿の共同開催により、北海道エリアのメガソーラ発電所（15か所）の合計出力を前日18時時点で精度よく予測する課題の下で「太陽光発電量予測技術コンテスト」が催された⁽⁶⁾。本稿は、同コンテスト参加時に適用した太陽光発電量の予測技法を概説した上で、予測誤差の分析結果等を概説する。

2. 予測技法の概要

まず太陽光発電量予測の使用データ、予測モデルの全体像を図1及び表1に示す。予測モデルの処理ステップは、大きく3ステップに分かれており（1）気象状態の予測、（2）日射強度の補正、（3）太陽光

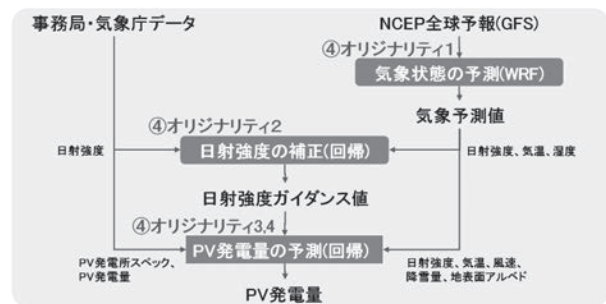


図1 太陽光発電量の予測ステップ

表1 予測ステップの全体像、使用データの詳細

①全体像 ステップ	②使用データの詳細				
	学習時の使用データ	予測時の使用データ	学習対象期間	データ入手先	有償/無償
気象状態の予測	・なし	・全球予報GFS(0.5度格子、UTC 0時)	なし	・米国NCEP	無償
日射強度の補正	・気象量予測値(各PVサイト) ・日射強度観測値(最寄地点)	・日射強度を含む気象予測値 ・学習済み回帰パラメータ	2016~2017年	・事務局 ・気象庁	無償
PV発電量の予測	・気象量予測値(各PVサイト) ・PVスペック(座標、容量、配置角) ・PV発電量実績値(各エリア)	・気象量予測値(各PVサイト) ・学習済み回帰パラメータ	2016~2017年	・事務局	無償

* 株式会社 東芝 研究開発センター

(以下 PV) 発電量の予測から構成される。

使用データの詳細は、全体の入出力データを学習時と予測時に分けて次のとおりである。まず学習時は、入力データがコンテスト事務局より提供されたメガソーラ発電所情報（緯度・経度、定格容量、平均レイアウト角）、過去2年間（2016年1月～2017年12月）の発電実績データ、気象予測データ、同時期の日射量観測データ（札幌、帯広、網走、室蘭の4気象官署）となり、出力データが各学習ステップで求めた学習済みパラメータ表（回帰係数）となる。次に予測時は、入力データが前述の学習済みパラメータ表、気象予測データとなり、出力データが太陽光発電量の予測値となる。なお気象予測データは、米国 NCEP から取得した全球予報 GFS（0.5 度）である。実行環境は、標準的な PC サーバ（Xeon, 64bit 版 Linux）である。計算時間は、気象状態の予測ステップが約2時間/日、日射強度の補正ステップ・PV 発電量の予測ステップが学習時5分以内、予測時1分以内となる。

3. 気象状態の予測

気象状態の予測ステップは、気象モデル WRF（Weather Research and Forecasting）を用いた全球予報 GFS の力学的ダウンスケーリング処理である。気象モデル WRF は、米国大気研究センター（NCAR: National Center for Atmospheric Research）の開発する領域気象シミュレータでありオープンソース化されている。この WRF は、初期値・境界値となる全球予報 GFS（約 50km 格子）と日本周辺の地形情報を入力に、大気の運動をシミュレーション計算した結果、解像度の高い気象予測値（9km 格子、30 分間隔）を出力する。WRF の出力ファイルから各メガソーラ発電所の座標を含む格子点値を抽出した上で、太陽光発電量予測に用いる。

全球予報 GFS は、毎日 UTC00 時（JST09 時）を初期化時刻とするデータを用いた。この初期化時刻は、GFS 公開時刻（初期化時刻の約3時間後）及び予測ステップ処理時間（約2時間/日）を加えた上で、毎日 JST18 時のデッドライン制約を十分に満たせるタイミングとした選択である。また気象モデル WRF のスピニングアップ時間は、初期化時刻 JST00 時から翌日の日出時刻 JST06 時頃までの約21時間を確保した。

気象予測データに関しては、気象庁及び民間気象会社から購入できるが、内製すると多数のメリットが考えられる。たとえば用途に応じて高精度化でき

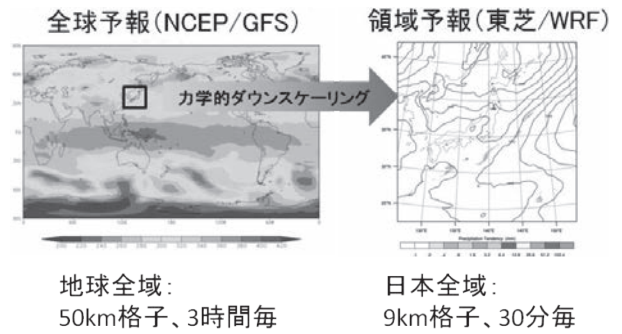


図2 気象予測モデルの運用

ることがメリットである。入力データが同じ全球予報である以上、質的予報（例：雨天、晴天）で精度差が出なくても、量的予報（例：日射強度、気温）で精度差を出せることがある。気象モデル WRF は、研究用シミュレータゆえに計算メッシュの切り方、放射伝達計算の方法など複数のパラメータを含んでおり、これらパラメータを最適化することで日射強度等の予測精度を向上できる。また機械学習と組み合わせることで、豊富な気象要素を説明変数に利用できることもメリットである。たとえば電力需要等の予測精度を高めたい場合、同じ気象要素を用いて回帰モデルの改良だけで精度を改善するには限界がある。気象要素のバリエーションを増やしてスパースモデリング技法で回帰モデルを改良することで、先ほどの限界を突破できることがある。

4. 日射強度の補正

日射強度の補正ステップは、気象予測値から補正済み日射強度（ガイダンス値）を作成する処理である。一般にガイダンス値とは、数値気象モデルの予測値を対応する観測値で校正して、数値気象モデルに含まれる系統的誤差を除いた予測値等である。報告者らの太陽光発電量の予測ステップは、各気象要素を発電量への影響が明らかな形で含むグレイボックス方式を採用しており、最も重要な気象要素として日射強度を含んでいる。そこで日射強度に限りガイダンス値を作成した。

ガイダンス値の計算式は、気象予測値を説明変数として日射強度の観測値を目的変数とする重回帰式である。教師データは日射強度の観測値（主催者の提供値、気象官署の観測値）、説明変数は日射強度の当該格子点を含む周辺格子点の統計値を用いた。また各メガソーラ発電所のガイダンス値を求める場合、直線距離で最寄りの地上観測点から決めた回帰係数を採用した。

図3は、あるメガソーラ発電所の座標で抽出した

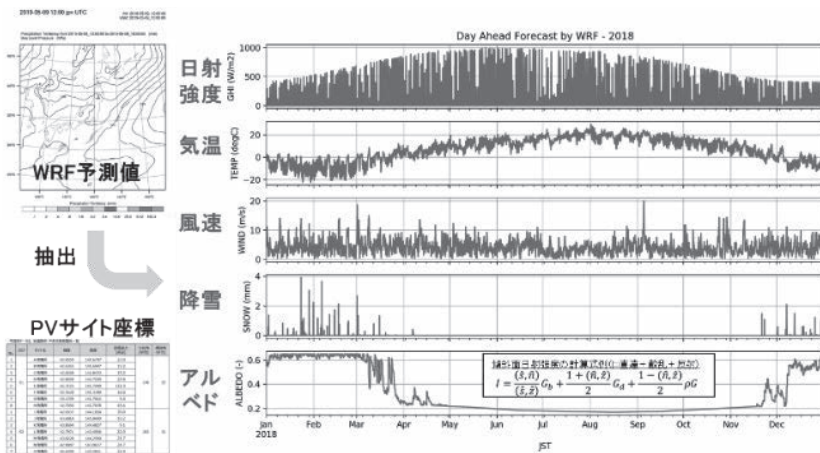


図3 気象モデルの出力例

(ガイダンス値を含む) 気象予測値を全年 (2018年1月～12月) で可視化した結果である。日射強度及び気温は、太陽光発電量の予測に不可欠な気象要素である。このほか発電モジュールの温度に影響する風速、発電モジュールの着雪に影響する降雪量、傾斜面日射強度に影響するアルベド(地表の反射率)の各気象要素を抽出した上で、太陽光発電量の予測に用いた。なお、太陽光発電量に対する風速とアルベドの効果は、それぞれ発電モジュール温度を推定する経験式、傾斜面日射強度を推定する経験式に基づき考慮しており、調整可能なパラメータを含まない。

5. PV 発電量の予測

PV 発電量の予測ステップは、ホワイトボックス方式のPV工学モデル(物理モデル)と機械学習(スパース回帰)を組み合わせるグレイボックス方式である。PV工学モデルは、「気象状況からモジュール傾斜面の日射強度を計算して、さらに気温及び風速から推定されるモジュール温度を考慮して換効率を計算して、最後にモジュール容量、傾斜面日射強

度、変換効率及びインバータ容量を考慮して発電量を計算する」方式である。PV工学モデルは、各気象要素の発電量への影響が明確で説明性の高さがそのメリットとなるが、パラメータ数が少なく発電量実績データへ合わせこむ自由度の低さがそのデメリットとなる。これらのメリットを維持してデメリットを解消するため、メガソーラ発電所の設備スペックを不定パラメータと捉えた上で、発電実績

データへの当てはまりが良くなるようにスパース回帰で推定する方式を考案した。

図4は、グレイボックス方式の概要と各ステップで予測誤差を評価した結果である。予測誤差は前日18時時点の予測値から実績値を差し引いた数値とし、rRMSE(%)は予測誤差(MW)の二乗平均平方根を平均出力(MW)で規格化した数値である。ホワイトボックス方式のPV工学モデルは、各メガソーラ発電所の定格容量と平均レイアウト角から合計発電量を予測する方式である。グレイボックス方式のLasso/Ridge回帰は、各メガソーラ発電所のレイアウト角と過積載率(=モジュール容量÷インバータ容量)をパラメータと見なして、複数の予測値の候補群をPV工学モデルから求めた上で、これら予測値の候補群とエリア合計実績値の関係をスパース回帰(正則化項付きの最小二乗法)することで、PV設備パラメータ推定と回帰式決定を同時に行い予測する方式である。グレイボックス方式(Lasso/Ridge回帰)とホワイトボックス方式(PV工学モデル)の予測誤差における相違は、レイアウト角と過積載率をパラメータと見なして最適化した

効果と解釈できる。なおLasso回帰はRidge回帰と比べて予測誤差が小さいが、正則化項の相違から係数ゼロが出やすい。係数ゼロは「発電所により合計出力に影響なし」と解釈できるので、物理的な実態と異なる。そこでLasso回帰とRidge回帰の予測値を併用した。なおスパース回帰の結果、推定されたレイアウト角や過積載率の設備パラメータは、解釈可能な数値であ

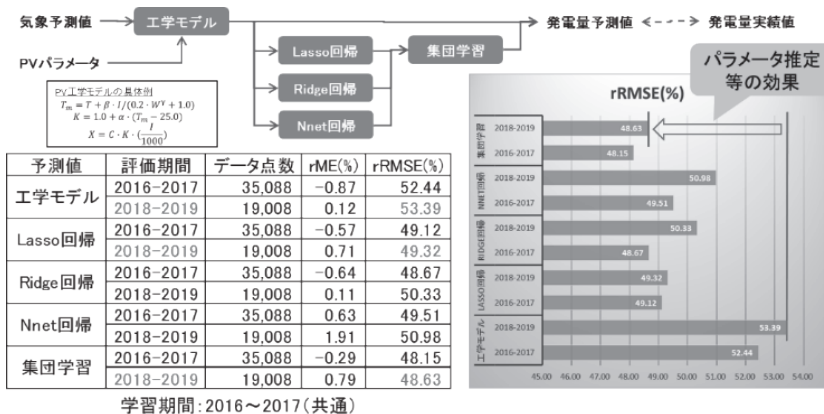


図4 太陽光発電量の予測モデル比較

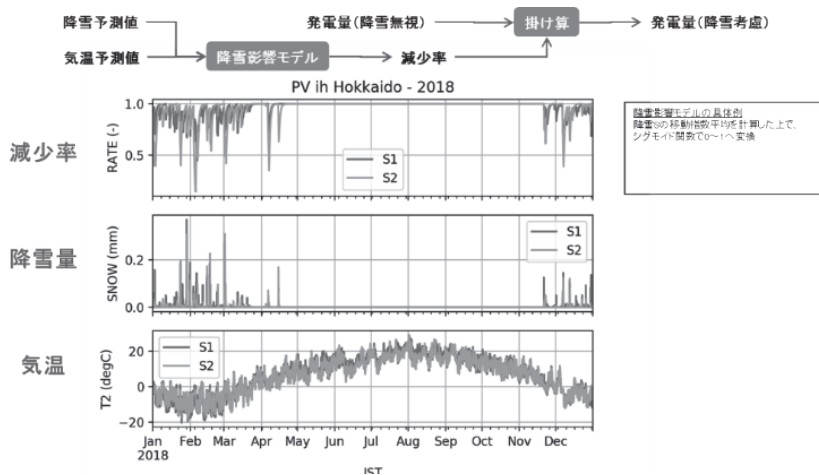


図5 降雪影響の簡易モデル

るが、正解データを保有しておらず有意性や推定誤差等の検証は未実施である。

最後の Nnet 回帰は、集団学習で精度を高める目的で導入したブラックボックス方式であり、スパース回帰と同じ入出力データ間の関係をニューラルネットで回帰する方式である。前記 Lasso/Ridge/Nnet 回帰の各予測値を集団学習で組み合わせた結果を最終的な予測値とした。集団学習は、各予測値の加重平均を考えて重み係数を Lasso 回帰で学習する方式とした。なお各予測方式は、予測誤差の小さい順に集団学習、Lasso 回帰、Ridge 回帰、Nnet 回帰、工学モデルとなるが、同じ気象予測値を入力するので出力の相違は少ない。

太陽光発電量への降雪影響は、簡易モデルを作成して予測の最終ステップで考慮した。まず太陽光発電量の実績値と気象予測値を並べて可視化すると、

冬期（12月～3月）の降雪時に発電量が少ない現象、降雪後に発電量の低下が数日間続く現象を観察できる。発電量の低下から回復までを降雪影響の事象1件と数えると、事象件数は学習期間の2年間に10件程度しかなく、学習と検証に十分な件数とは言えない。そこで降雪影響は、図5の様に降雪量予測値の移動指数平均から約3日影響の残る変数を作成した上で、シグモイド関数で [0,1] 変換しただけの発電量減少率を作成した。

これら簡易モデルのパラメータ3個は、予測誤差を最小化するグリッドサーチから決定した。降雪量を考慮した発電量予測値は、集団学習方式の発電量予測値に減少率を乗算した結果とした。

6. 予測結果

図6は、最終的な太陽光発電量の予測値と実績値との可視化例である。なお予測誤差は、可視化時に分かりやすい定格容量による規格化を用いており、先ほどと規格化が異なる。両者の関係は、太陽光発電の稼働率を約13%と考えると「定格出力で規格化した予測誤差 \div 平均出力で規格化した予測誤差 \times 0.13」である。図6左側のヒストグラム (N = 55週) は、週毎の予測誤差を定格容量で規格化した相対 RMSE 値の分布である。通年の平均値は相対 RMSE が約6% (相対 MAE が約3%) となる。相対 RMSE 値に関して最小の週 (2018/3/5 ~ 3/11)、平均的な週 (2018/10/22 ~ 10/28)、最大の週 (2018/12/10 ~ 12/16) を図6に示した。誤差最大の週を除くと、概ね良好な前日予測と確認できる。誤差最大の週は、降雪影響の考慮不足が原因で予測値が実績値を大きく上回っている。

図6のヒストグラムで誤差の大きい週 (相対 RMSE \geq 8.0%) は5週あり、うち3週が冬期で同様の傾向が見られる。これらの大きな誤差の原因は、発電モジュールが雪に覆われる状況を

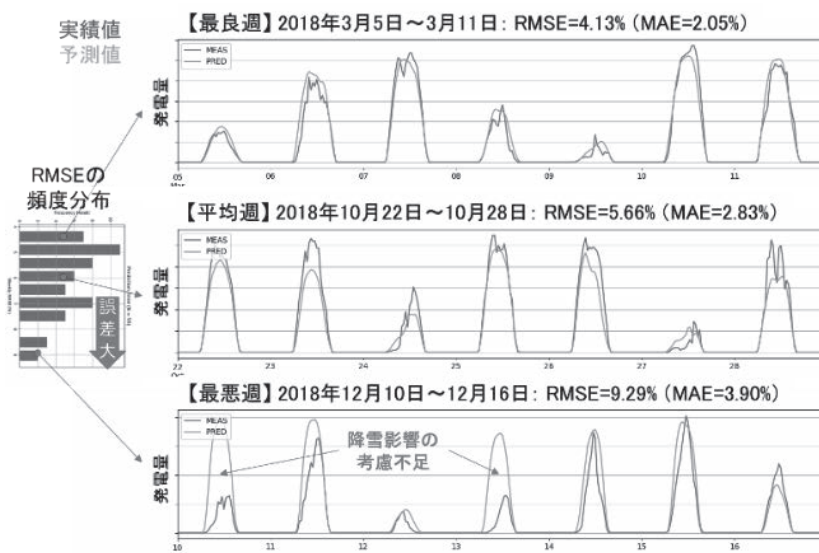


図6 太陽光発電量の予測値と実績値

予測時に正しく考慮できていないことにある。このように雪国の太陽光発電量を予測するには、降雪量の予測だけでは不十分であり、発電モジュール上の積雪状態を予測時に考慮する必要がある。

7. おわりに

太陽光発電量の前日予測技術として、予測精度の高さと説明性の高さを両立する東芝のグレイボックス方式を概説した。この方式は、気象モデル WRF で予測した高解像度かつ多様な気象要素を用いて予測精度を高める点、PV 工学モデル（物理モデル）とスパース回帰を組み合わせて説明性と予測精度を両立できる点に特徴がある。今回のコンテスト参加に際して、予測精度を高める集団学習、太陽光発電への降雪影響予測モデルを追加した。

予測精度に関しては、グレイボックス方式だけでも良好だが、ニューラルネット回帰のブラックボックス方式を集団学習で加えると、わずかな改善が見られた。ただし各予測方式は、同じ気象予測を用いるので結果の多様性が少ない。またグレイボックス方式で推定した PV 設備パラメータ（レイアウト角、過積載率）は、解釈可能な数値であるが、正解データがなく推定誤差等を検証できなかった。降雪影響モデルに関しては、学習期間中の事象数も少なく不十分な結果となった。降雪影響モデルは、建築分野等に関連する研究が存在するため、これらを参考に物理的メカニズムを明らかにした上で、学習を組み合わせるグレイボックス方式が有効と考えられる。本稿が、将来における太陽光発電量予測技術の発展に少しでも役立てば幸いである。

8. 参考文献

- 1) “System Integration of Renewables - An update on Best Practice,” International Energy Agency, January 2018 (NEDO 翻訳有)。
- 2) L. Gigoni et al., “Day-Ahead Hourly Forecasting of Power Generation From Photovoltaic Plants,” in IEEE Transactions on Sustainable Energy, vol. 9, no. 2, pp. 831-842, April 2018.
- 3) B. Carlo et al., “The value of day-ahead solar power forecasting improvement” in Solar Energy, vol. 129, pp. 192-203, May 2016.
- 4) T. Hong et al., “Probabilistic energy forecasting : Global Energy Forecasting Competition 2014 and beyond,” in International Journal of Forecasting, vol.32, March 2016.
- 5) 加藤丈佳, 予測コンペの開催報告, 平成 28 年度 電気学会 C 部門大会講演論文集 TC8-5, pp. 234-239, 電気学会 (2016-08-31)。
- 6) 太陽光発電量予測技術コンテスト「PV in HOKKAIDO」の結果について, 東京電力ホールディングス株式会社・北海道電力株式会社プレスリリース (2019/06/28)。

著者略歴



進博正 (シン ヒロマサ)

1999 年 3 月京都大学大学院理学研究科物理学・宇宙物理学専攻修了。
同年 4 月株式会社東芝入社。研究開発センター勤務。

主にスマートグリッド関連技術の研究開発に従事。気象予報士。



志賀慶明 (シガ ヨシアキ)

2016 年 3 月立教大学理学研究科物理学専攻博士後期課程単位取得退学。
同年 4 月株式会社東芝入社。研究開発センター勤務。

2017 年 3 月立教大学理学研究科物理学専攻博士後期課程、博士 (理学) 取得。

主に、電力分野の予測技術の研究に携わる。



柿元満 (カキモト ミツル)

1982 年 3 月東京大学理学部物理学科卒業。

1987 年 3 月東京大学理学系大学院物理学専門課程卒業。

同年 4 月株式会社東芝入社。計算機システム開発に従事。

現在研究開発センターシステム技術ラボラトリー勤務。

主に、数値気象シミュレーションを利用した太陽光発電量予測の研究に携わる。理学博士。