

# Just-In-Time Modeling に基づく日射量予測手法のファジィクラスター分析による改良

## Improvement of Just-In-Time Based Solar Radiation Prediction Method by Using Fuzzy Clustering

立命館大学 ○福原 大祐, 江田 優矢, 竹田 皓貴, 鷹羽 浄嗣, 吉川 直樹, 天野 耕二

Daisuke Fukuhara, Yuya Eda, Koki Takeda, Kiyotsugu Takaba, Naoki Yoshikawa and Koji Amano  
Ritsumeikan University

**Abstract**—This paper proposed the improved method of Just-In-Time based solar radiation prediction by using fuzzy clustering. We used black-box model (JIT Modeling) with the physical model (GPV data) for solar radiation prediction method. In JIT modeling, there is a procedure to search for similar data. We improved searching for similar data procedure with fuzzy cluster. The search speed of similar data increased by proposed procedure. We clarified that prediction accuracy can be improved by optimizing fuzzy cluster parameters.

**Key Words:** Solar radiation prediction, Fuzzy cluster analysis, Just-In-Time Modeling

### 1 はじめに

近年、環境問題やエネルギーセキュリティの観点から再生可能エネルギーへの関心は高まっている。そして再生可能エネルギーの中でも、設置が容易な太陽光発電は大量導入が期待されている。しかしながら、太陽光発電は季節、時間、気象条件によって発電量が変動するため、太陽光発電単体での安定的な電力供給は困難である。太陽光発電を安定的かつ効率的に電力供給するために、既存の系統電力やマイクログリッドのような電力網に接続することが考えられる。その際に、太陽光発電の不安定性を抑えるために事前に太陽光発電の発電電力量を予測し、発電電力量の予測情報を基に電力需給制御することが必要となる。

太陽光発電の発電電力量の予測に必要な日射量を予測する手法として、鈴木ら<sup>1)</sup>によって提案された、気象予報データを用いたデータベース型ブラックボックスモデリングの一種である Just-In-Time(JIT) Modeling が有効であることが知られている。

また、筆者ら<sup>2)</sup>の先行研究では、JIT Modeling で予測された日射量に対して、予測誤差を時間相関があるものと仮定した AR モデルによって補正することで、予測精度を向上させている。

JIT Modeling では予測対象の入力データと類似データを、データベース内の全てと比較してから選択する必要がある。本稿では、JIT Modeling をファジィクラスター分析の代表的な手法である Fuzzy c-means(FCM)法を用いてデータベースを事前にファジィクラスター分析することで、データベースの探索範囲を絞り予測精度を向上する方法を提案する。本稿では、通常の JIT Modeling と FCM 法を用いて改良した JIT Modeling について、単一地点の気象予報データを用いて予測精度の比較を行った。

### 2 日射量予測アルゴリズム

#### 2.1 Just-In-Time(JIT) Modeling

本稿では、ブラックボックスモデリングの一種である Just-In-Time(JIT) Modeling を用いて、日射量の予測を行った。JIT Modeling では、過去の気象データよりデータベースを構築し、予測対象の気象データを入力データとして、データベースと比較し類似している近傍のデータを抽出する。抽出した近傍のデータから局所モデルを構築して、予測値を算出する。

#### 2.2 データベース

2011年から2016年の気象庁のMSM-GPV データ<sup>3) 4)</sup> および日射量データ<sup>5)</sup> をデータベースとして利用した。MSM-GPVは、日本の気象を5kmごとに39時間後までのシミュレーションを行う数値予測モデルである。内容は、気圧、風(南北/東西)、気温、湿度、時間降水量、雲量(高層/中層/低層)があり、1日に8回更新されている。

上記の9つの気象情報に日射量に影響すると思われる太陽高度および可照時間の理論値を加えた11データの0時における1日分(0時から23時)を入力データ  $x^j(h)$ , ( $h=0, \dots, 23$ )とし、水平面日射量の実測値1日分(0時から23時)を出力値  $y^j$  として  $j$  番目のデータを構成する。

#### 2.3 類似データ選択

本稿では、日射量を目的変数、気象データを説明変数とし、重回帰分析の偏回帰係数を求め、重みとして利用した。また、最適な説明変数を赤池情報量基準(AIC)<sup>6)</sup>に基づく変数減少法により決定<sup>7)</sup>し、その係数を利用した。

このとき、入力データとデータベースの類似度の指

標には(1)式に示す重み付きユークリッド距離を用いた。

$$d(q, j) = \sqrt{(x^q - x^j)^T W_p (x^q - x^j)} \quad (1)$$

ここで、 $x^q$ :予測対象の入力データ(ベクトル)、 $x^j$ :データベースの $j$ 番目の入力データ(ベクトル)、 $p$ :入力次元数、 $W^p$ : $p \times p$ の重み行列とする。

近傍データの選択方法は、K-Nearest Neighbors (K-NN)を用いた。K-NNは、予測対象の入力ベクトルとデータベース内の入力ベクトル群の算出した距離 $d$ の中から、小さい順に $KNN$ 個を選択する方法である。本稿では、 $KNN=20$ と設定した。

## 2.4 予測値の算出

抽出した近傍データから、日射量の予測値を(2)式に示す重み付き線形平均法を用いて算出した。

$$y^q = \frac{(\sum_{j=1}^{KNN} \omega_j y^j)}{(\sum_{j=1}^{KNN} \omega_j)} \quad (2)$$

ここで、 $y^q$ :予測対象の出力値、 $y^j$ :近傍の出力値、 $\omega_j$ :近傍に対する重みとする。

(2)式の $\omega_j$ は、重み関数として(3)式に示すTricube関数を用いて算出する。

$$\omega_j = \left(1 - \left(d_j/d_k\right)^3\right)^3 \quad (3)$$

ここで、 $d_j$ :予測対象と近傍の距離、 $d_k$ :予測対象から最も遠い近傍との距離とする。

## 2.5 Fuzzy c-means(FCM)法

通常のクラスター分析では、各データは1つのクラスターに所属するが、ファジィクラスター分析では複数のクラスターへ所属できる。気象予報データを通常のクラスター分析で分類すると、小さいクラスターになり十分なデータ数が集まらない場合があるが、ファジィクラスター分析では2つ以上のクラスターの間へ所属できることを利用して、より少ないクラスター数で分類できるため、ファジィクラスター分析を行った。

本稿ではファジィクラスター分析の代表的な手法であるFuzzy c-Means(FCM)法<sup>7)</sup>を用いて、クラスター分析を行った。FCM法では各代表点 $u_k$ および各クラスター $k$ への帰属度 $u_{ki}$ (メンバーシップ)で、データ $x_i$ を $c$ 個のクラスターに分類して表現される。最初に帰属度 $u_{ki}$ をランダムに初期化し、帰属度 $u_{ki}$ より各クラスター $k$ の重心を(4)式より算出し代表点 $u_k$ とする。次に、各代表点 $u_k$ と対象データより(5)式を用いて帰属度 $u_{ki}$ を更新する。帰属度 $u_{ki}$ と代表点 $u_k$ が収束するまで反復を続けることでファジィクラスター分析の結果を得ることができる。ただし、 $m$ は $m>1$ でありファジィの度合いを決めるパラメータである。

$$u_k = \frac{\sum_i^N (u_{ki})^m x_i}{\sum_i^N (u_{ki})^m} \quad (4)$$

$$u_{ki} = \left[ \sum_j^c \left( \frac{\|x_i - u_k\|}{\|x_i - u_j\|} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (5)$$

$ki$ は各データの各クラスターへの帰属度 $u_{ki}$ を指定する変数である。 $m$ については、データが各クラスターへの帰属度合が変化することから、 $m$ が小さいほど各クラスター間の帰属度の差が大きくなる。

本稿では、2.2節に記したデータベースを標準化し、主成分分析により次元削減した上でファジィクラスター分析により分類することとした。事前にデータベースをクラスター分類することで、予測対象の入力データの所属クラスターから、類似データの選択の探索範囲を絞ることができる。

各クラスター $k$ の帰属度が任意の閾値を上回ったデータ群を $X_{uk}$ とする。予測対象の入力データの帰属度が任意の閾値を上回った場合には、データベース内の $X_{uk}$ から類似データを選択する。予測対象の入力データの帰属度が任意の閾値を下回った場合には、次の2通りの方法を用いた。まず、(6)式に示す積集合 $O_u$ を求める。

$$O_u = \bigcap_{k=1}^c \overline{X_{uk}} \quad (6)$$

ここで $X_{uk}$ の補集合を $\overline{X_{uk}}$ とし、(a)(b)いずれかの方法で出力値を得た。

- データベース内の和集合 $X_{uk} \cup O_u$ から類似データを選択する。選択した類似データから出力値を得る。
- データベース内の和集合 $X_{uk} \cup O_u$ から類似データを各クラスターで選択する。各クラスター $k$ で選択したデータから得た出力値を、予測対象の各クラスター $k$ への帰属度で加重平均を取り出力値とする。

## 2.6 太陽光発電の発電電力量

太陽光発電の発電電力量 $p_r$ は、実測の日射量または出力された日射量の予測値を用いて(7)式より算出する。

$$p_r = PR \cdot W_{nmo} \cdot y_{pv} \quad (y_{pv} = \alpha \cdot y^q) \quad (7)$$

ここで $W_{nmo}$ は公称最大出力、 $PR$ (Performance Ratio)は太陽光発電の性能比、 $\alpha$ は水平面日射量を斜面日射量 $y_{pv}$ に変換する係数である。

## 2.7 予測精度の評価手法

予測精度の評価として、(7)式に示すRMSE(Root Mean Square Error)を用いた。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (mv_i - pv_i)^2} \quad (7)$$

ここで、 $N$ :データ数、 $mv$ :実測値、 $pv$ :予測値とする。比較には、大阪の単一地点における2016年の気象予報データを用いた。

### 3 結果と考察

Fig.1 と Fig.2 に通常の JIT Modeling による予測結果の RMSE と JIT Modeling に 2.5 節で示した手順(a)(b) (以下, それぞれ Fuzzy(a), Fuzzy(b)と表す) を適応した結果による RMSE を示す.

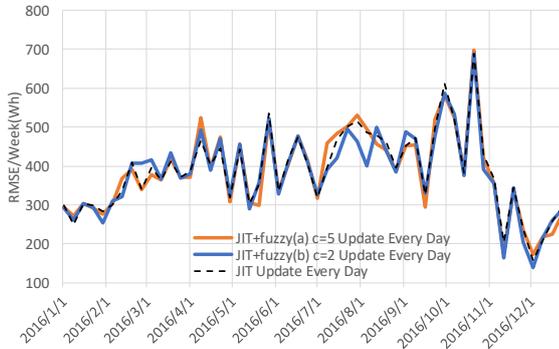


Fig. 1: RMSE on updating database every day.

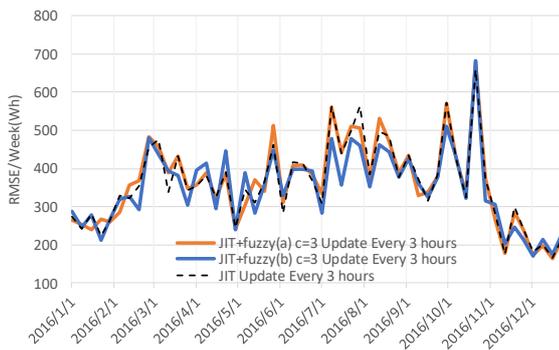


Fig. 2: RMSE on updating database every 3 hours.

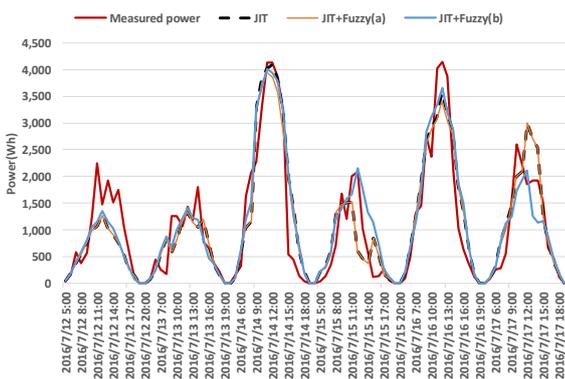


Fig. 3: Power on updating database every 3 hours.

Fig.1 は, 毎日 0 時に 24 時間分の気象データを更新するケースにおいて, 1 週間単位の RMSE を 52 週分示したものである.点線が JIT Modeling のみによる太陽光発電の予測の RMSE である.実線がそれぞれ JIT Modeling にクラスター数  $c=5$  として Fuzzy(a),  $c=2$  して Fuzzy(b)を加えたものである. Fuzzy(b)の実線の

RMSE が点線よりも特に下回る週は 8 月初めの週であり, 80Wh 程度精度が改善されている.2 月末の週は, Fuzzy(b)が点線を上回り 60Wh 程度精度が下がっている.その他の多くの週は点線に実線に重なっている.

Fig.2 は, 3 時間毎に気象データを更新した際の, クラスタ数  $c=3$  として 1 週間単位で RMSE52 週分を示したものである.JIT Modeling に Fuzzy(b)を加えると, 精度向上した週と変化がない週がほとんどであり, 精度が下がっている週が 24 時間毎に更新したときよりも少なくなっている.このことから特に気象データの更新が 3 時間の際には, JIT Modeling に Fuzzy(b)を組み合わせると精度向上には役立つと言える.また, RMSE の変化が無い箇所も精度向上には貢献しないものの, 類似データの探索範囲が絞ることができているため演算速度が改善されている.

Fig.3 は Fig.2 で精度向上が確認できた 7 月中旬で,

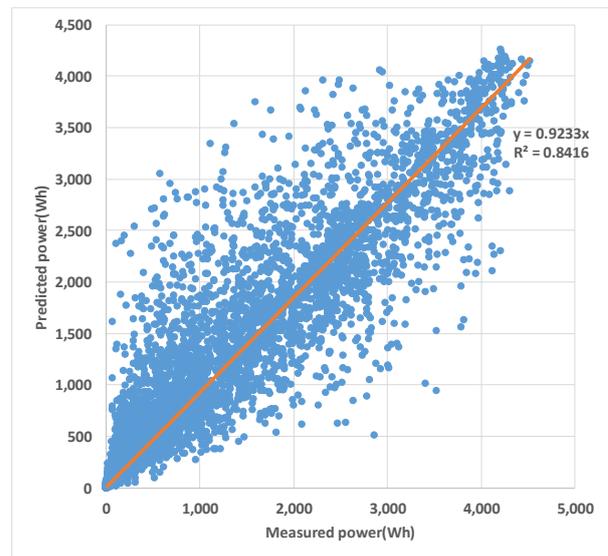


Fig. 4: JIT scat plot (updating database every 3 hours).

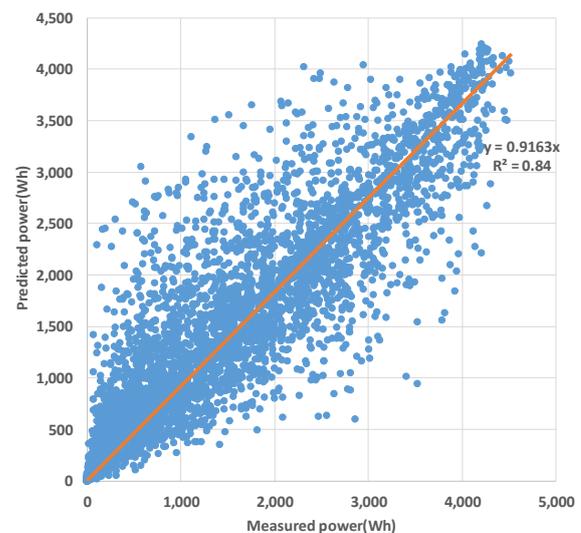


Fig. 5: JIT + Fuzzy (a) scat plot (updating database every 3 hours).

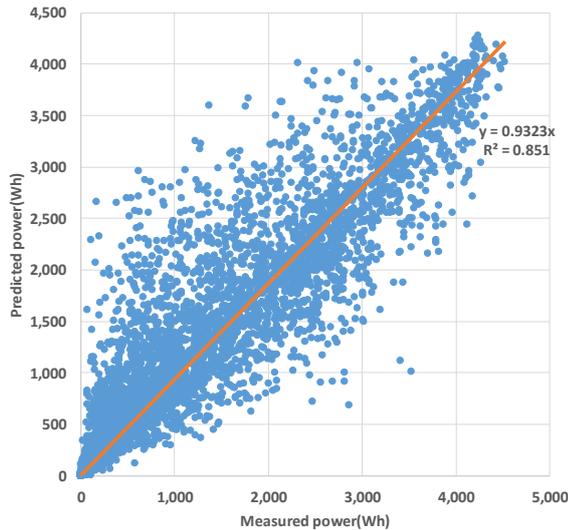


Fig. 6: JIT + Fuzzy (b) scat plot (updating database every 3 hours).

赤線の観測された日射量から算出した発電電力量と、日射量の予測値から算出した発電電力量を比較したものである。JIT Modeling に Fuzzy(a)を加えたモデルは、通常の JIT Modeling と同じ予測値を出力している。それに対して JIT Modeling に Fuzzy(b)を加えたモデルは、JIT Modeling の予測値が実測値から外れている 3 日間の内 2 日間で精度が向上している様子が明らかである。

Fig.4, Fig.5, Fig.6 は、観測された日射量による発電電力量と、3 時間毎に気象データを更新した際に、各モデルによって出力された 2016 年の日射量予測値を基に算出した発電電力量の関係をプロットし、最小二乗法による切片 0 の回帰直線を示したものである。 $y = x$  に近いほどモデルの予測精度が高いと考えられ、3 時間毎に気象データを更新した際には Fuzzy(a)を加えたモデルでは予測精度が全体として下がっている。対して、Fuzzy(b)を加えたモデルは、JIT Modeling のみのモデルと比較して傾きが増加しており、モデル全体として予測精度が向上していることが見て取れる。これらのことから 3 時間毎に気象データを更新した際には、JIT に Fuzzy(b)を加えると予測精度が向上すると言える。

Fig.7, Fig.8, Fig.9, Fig.10 は、各モデルの 1 週間単位の RMSE52 週分を、クラスター数  $c=2\sim5$  の範囲で箱ひげ図で比較したものである。Fig.7, Fig.8 から JIT Modeling のみのモデルと比較して、クラスター数  $c=2$  の時に RMSE の四分位範囲 最小値と平均が下がっている。このことから 24 時間毎に気象データを更新する場合には、クラスター数  $c=5$  の時に Fuzzy(a)が、 $c=2$  の時に Fuzzy(b)の予測精度が高くなることが明らかとなった。

Fig.9, Fig.10 から JIT Modeling のみのモデルと比較して、クラスター数  $c=3$  の時に、JIT Modeling に

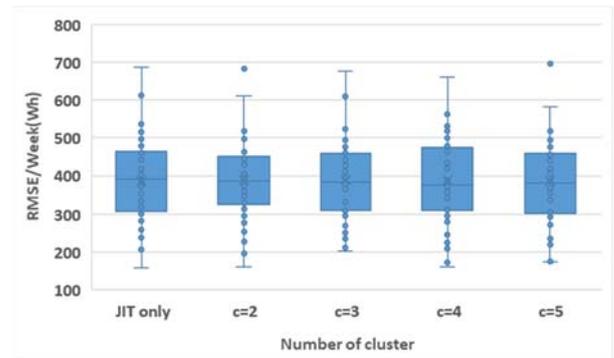


Fig. 7: JIT + Fuzzy (a) box plot (updating database every day).

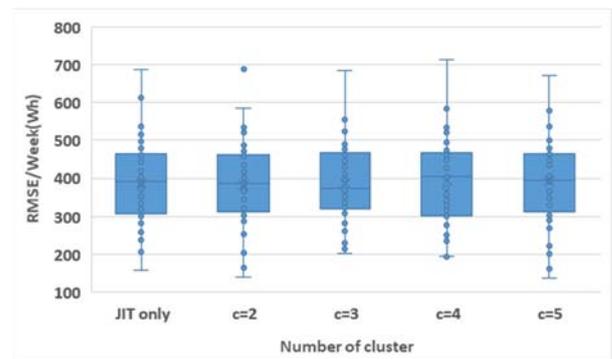


Fig. 8: JIT + Fuzzy (b) box plot (updating database every day).

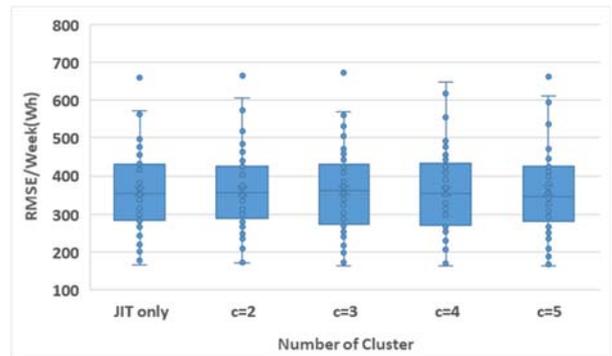


Fig. 9: JIT + Fuzzy (a) box plot (updating database every 3 hours).

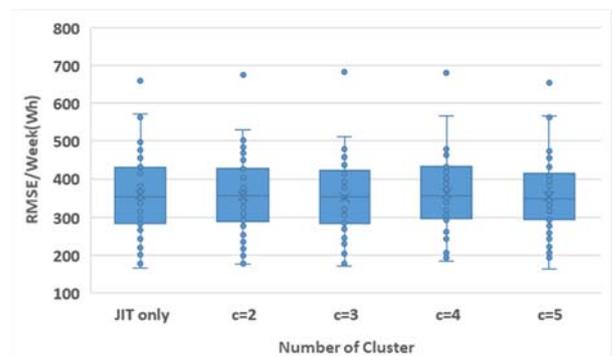


Fig. 10: JIT + Fuzzy (b) box plot (updating database every 3 hours).

Fuzzy(b)を加えたモデルの予測精度が RMSE の外れ値を除いた最大値と平均が下がっている.このことから 3 時間毎に気象データを更新する場合には, クラスタ数  $c=3$  の時に JIT Modeling に Fuzzy(b)を加えたモデルの予測精度が高くなることが明らかとなった.

24 時間毎に気象データを更新するとき, Fuzzy(a)はクラスタ数  $c=5$ , Fuzzy(b)は  $c=2$ , 3 時間毎に気象データを更新する場合にはクラスタ数  $c=3$  の時に, 予測精度が向上することから, クラスタ数  $c$  は予測に用いるデータベースによって, 最適化が必要であることが明らかである. また, Fuzzy(b)で 3 時間毎に気象データを更新する組み合わせが, 最も RMSE の外れ値を除いた最大値と平均が下がっていることから最も予測精度が向上する組み合わせであることが明らかとなった.

## 4 おわりに

本稿では, ファジィクラスタ分析を用いて JIT Modeling の改良を試みた. ファジィクラスタ分析によって予測精度が向上される場合と予測精度の向上に貢献しない場合とが確認できたが, いずれの場合も類似データの探索範囲を絞ることで, 演算速度の向上に貢献していることが確認できた. このことから, 予測精度を向上するためには, 気象データベースのファジィクラスタ分析においてパラメータを最適化することが必要であることが明らかとなった.

### 謝辞

難波巧氏, 船引伸哉氏には気象データの整備においてご協力いただいた. ここに記して謝意を表する.

### 参考文献

- [1] 鈴木, 後藤, 寺園, 若尾, 大関: Just-In-Time Modeling に基づく日射量予測手法の開発; 電気学会論文誌 B(電力・エネルギー部門誌), Vol.131, No.11, pp. 912-919 (2011)
- [2] 竹田, 鷹羽: PV 発電量予測に基づくマイクログリッドの確率的モデル予測制御; 計測自動制御学会論文集, Vol.54, No.2, pp. 219-226 (2018)
- [3] 一般財団法人気象業務支援センター-メソ数値予報モデル GPV(MSM); <http://www.jmbc.or.jp/jp/online/le/f-online10200.html>
- [4] 筑波大学 計算科学研究センター気象庁 GPV データ; <https://www.ccs.tsukuba.ac.jp/researchproject/database>
- [5] 国土交通省 気象庁-各種データ・資料; <http://www.jma.go.jp/jma/menu/menureport.html>
- [6] 赤池, 甘利, 北川他:赤池情報量規準 AIC—モデリング・予測・知識発見, 共立出版(2007).

- [7] Ross, Timothy: J. Fuzzy Logic with Engineering Applications, Third Edition, pp.352-353 (2010)