

AI(機械学習)による 日本卸電力取引所(JEPX)スポット市場の 市場価格予測

加藤 史葉

Weather Data Science

2018年7月

1. 電力市場とは

実は2000年に始まっていた電力自由化(特別高圧限定)。自由化対象が拡大していく過程の2005年に『日本卸売電力取引所(JEPX)』が開設され、現在もここで日々電気が売買されています。※電気の特性上キャピタルゲインは狙えない取引市場です

今回は、JEPX(以下、電力市場)の取引価格を予測するアルゴリズムを作りました。



JEPXには様々な市場がありますが
今回のターゲットは
スポット市場

D日に消費される電気は(D-1)日のスポット市場(10:00入札締切)で売買されるオペレーションを想定し
取引日の朝 = (D-1)日の朝に予測値をアウトプット

システムプライスではなく
九州のエリアプライス

30分毎48コマ/day → 43コの価格予測値

予測!!

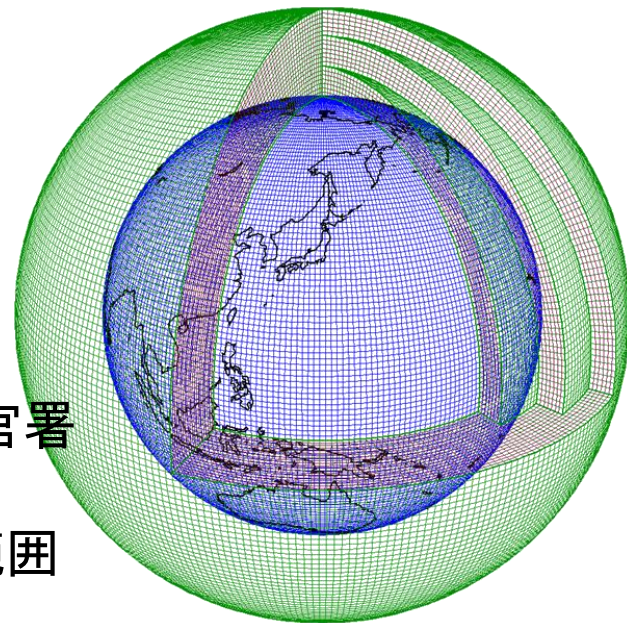
2. データの収集方法

市場価格データ

- JEPXのHPよりDL
- 2017年12月7日～

説明変数①: 気象庁MSM-GPV21zBASE

- 京都大学生存圏研究所データベースより取得
- データ切り出し対象地点は九州全7県の気象官署
- 日射量、気温、相対湿度、降水量
- MSM-GPV21zBASEは翌日21時までが予測範囲
- 2017年12月5日(12月7日予測対象)～



説明変数②: 国民の祝日等データ

- 内閣府HPよりDL
- 年末年始休暇の情報を追加

国土交通省 気象庁 Japan Meteorological Agency

キーワードを入力し検索ボタンを押下ください。 POWERED BY YANAGI

本文へ | ENGLISH | ご意見・ご感想 | サイトマップ

ホーム | 防災情報 | 各種データ・資料 | 知識・解説 | 気象庁について | 案内・申請

ホーム > 気象庁について > 報道発表資料 > 平成29年報道発表資料 > 日射量予測データの提供を開始します

日射量予測データの提供を開始します

報道発表日
平成29年12月5日

概要
スーパーコンピュータを活用した「数値予報」により計算された日射量予測データの提供を、本日(12月5日(火))より開始します。

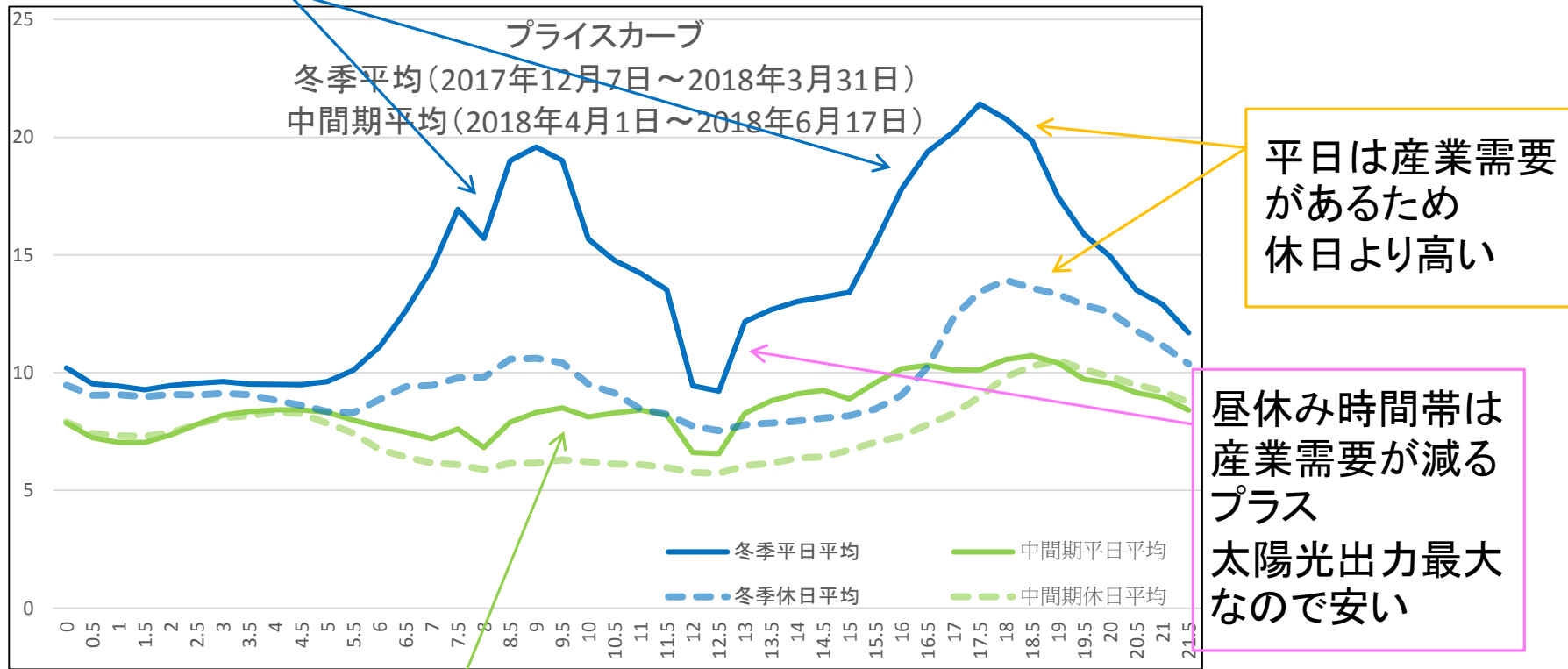
本文

3. プライスカーブを確認①『冬季は2コブ型、中間期は凹型』

早朝は需要の立ち上がり時間帯、

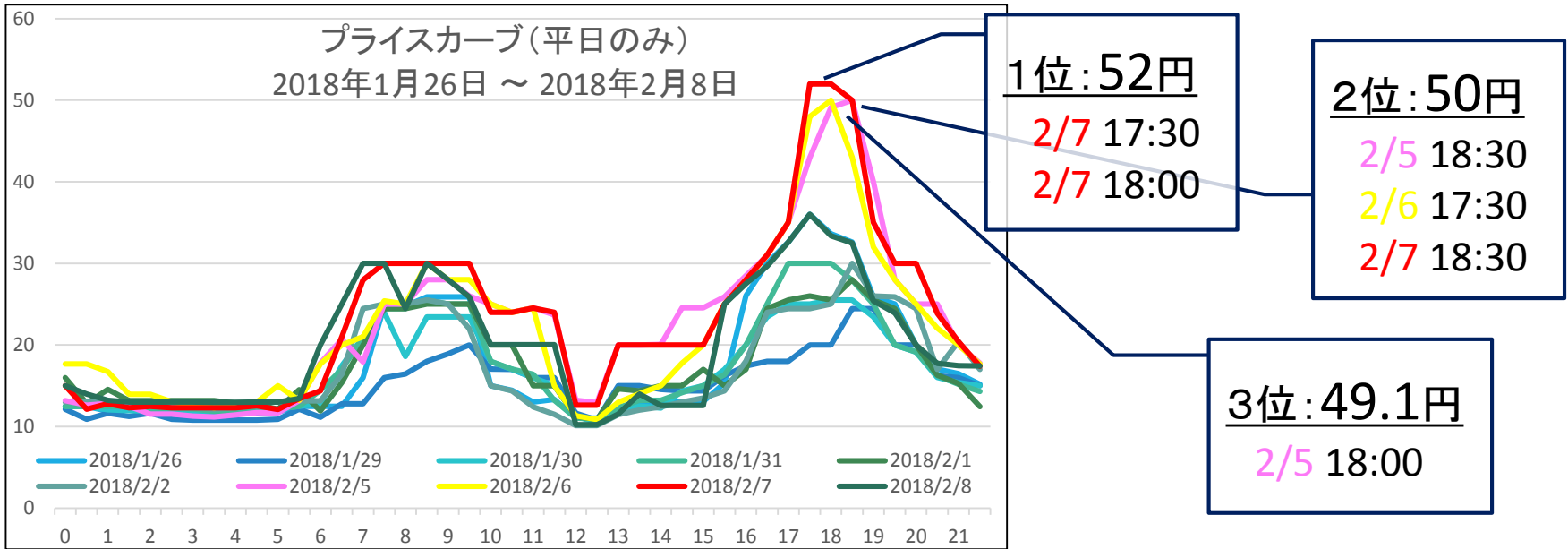
夕方はまだ産業需要が残りながら家庭需要も立ちあがってくる時間帯(需要量**多**)

早朝は太陽光出力フェードインの時間帯、夕方はフェードアウトの時間帯(供給量**少**)



- ・ 中間期は空調需要がないので冬季に比べて全体的に価格が安い
- ・ 春の中間期は年間で一番日射が強い時期と合致するので、日中は供給過多

3. プライスカーブを確認②『雪空厳寒日夕方に最高値』



livedoor NEWS

ニューストップ > 国内 > 社会 天気

5日(月)午前9時

九州に今世紀最強クラスの寒気が
最高気温が今季最低を更新する予想

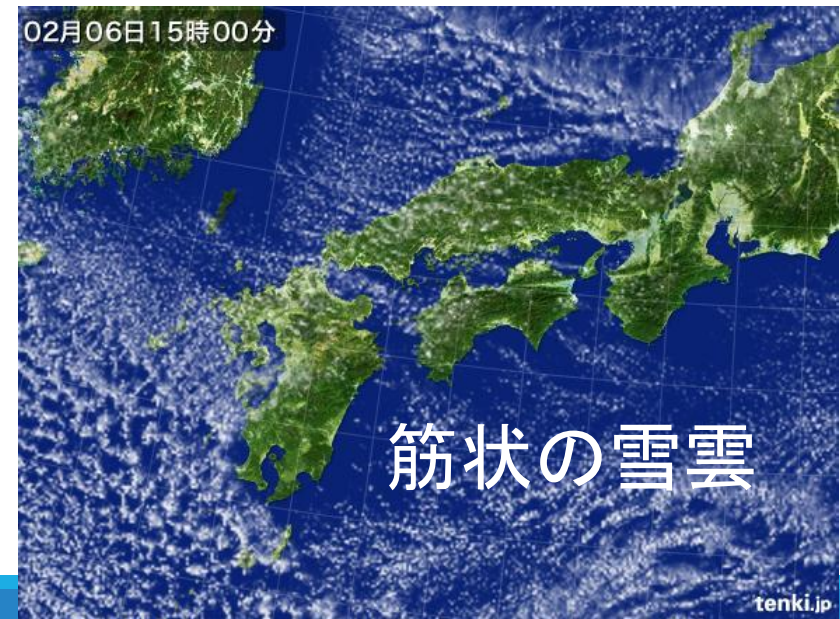
f 19 | 12 | 2018年2月5日 7時00分

上空の寒気
1500m付近 -6度
平地でも雪

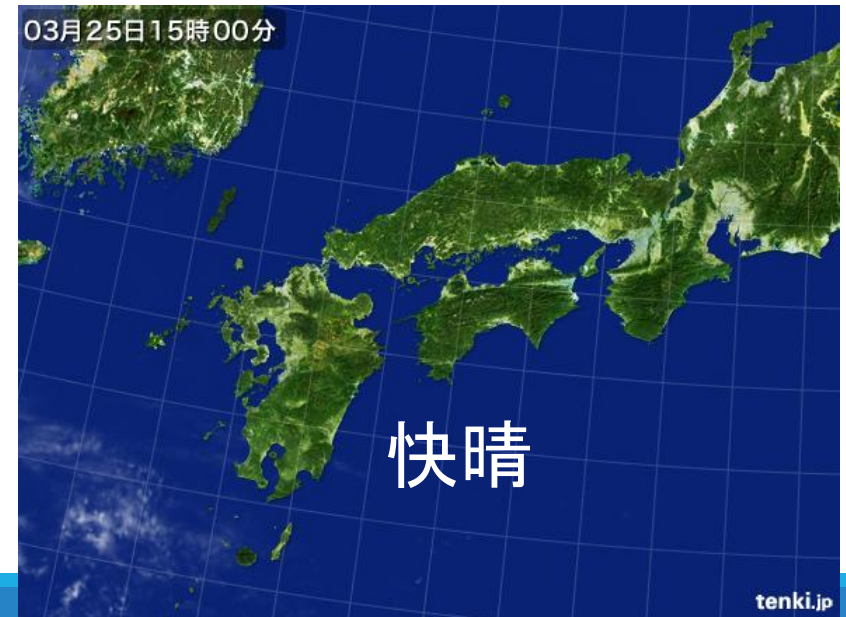
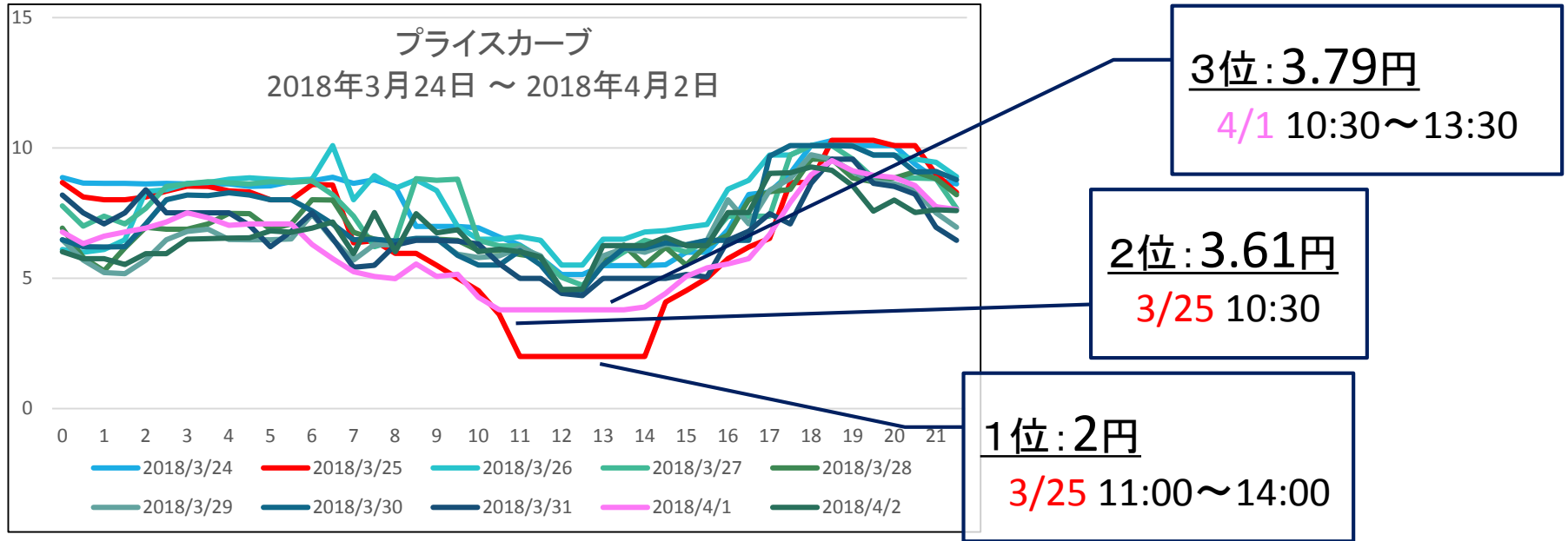
奄美(名瀬)

ざっくり言うと

- ✓ 九州の上空に、近年では珍しいほど強い寒気が流れ込む見込みだという
- ✓ 雪が降ることが少ない九州南部まで寒気に覆われている
- ✓ この影響により、日中の最高気温が今季最低を更新する可能性があるとのこと



3. プライスカーブを確認③『行楽日和の日曜日に最安値』



4. 初期仮説・分析アプローチ『価格は太陽光が握る』

仮説

- 九州エリアの市場は特に、売りに出てくる電気のうち太陽光の占める割合が大きい
ため、晴れれば供給過多で市場価格が下がる、曇雨天ならば上がる
 - ☞ 『**市場価格は太陽光が握っている？**』
- 市場価格は気温(空調需要)によって変動する(仮説というより既知のこと)

分析アプローチ

- コマ毎に学習させる
- 夏は気温が高いほど需要増、冬は気温が低いほど需要増(夏と冬で逆方向の相関)
 - ☞ 空調需要のない20℃を基準に気温との差分値を採用
- 一定期間の学習を日々ローリングしながら予測させる
 - ☞ たった半年分のデータしかないから

5. 1stアプローチ

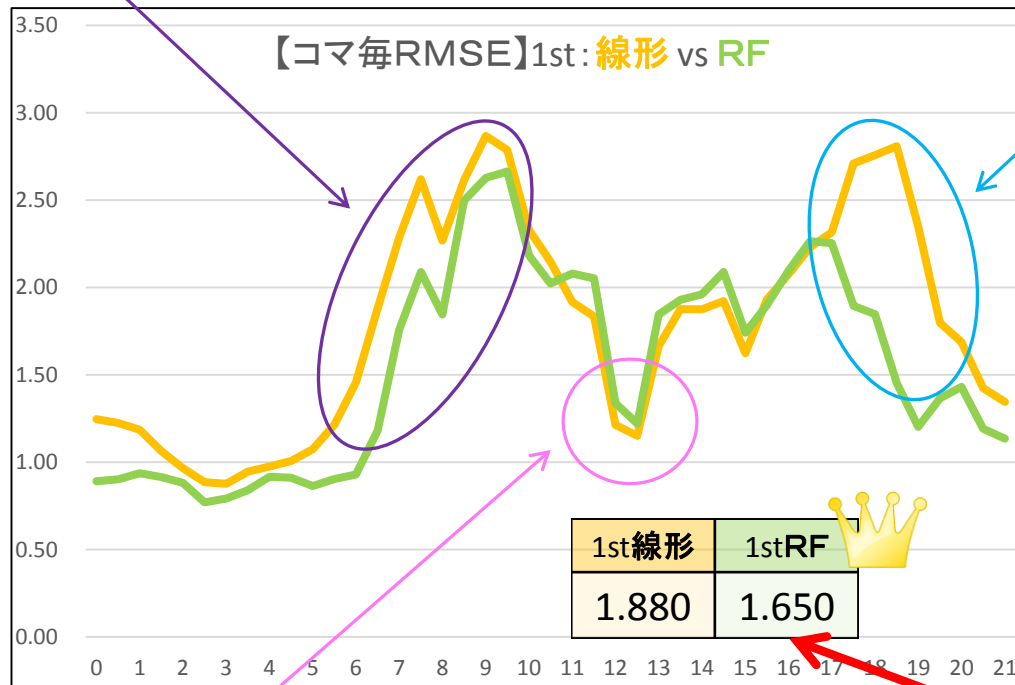
1stアプローチ	
特徴量	カレンダー情報(月:0~日:6 + 祝:7) 気象予測値(日射量・気温(差分値)・相対湿度・降水量) × 九州7県分
学習の流れ	コマ毎に学習 ウィンドウ期間過去3ヶ月(90日)をローリング学習
使用モデル	線形回帰、ランダムフォレスト(木の数=150/深さ=指定なし)
処置	気象予測値の多重共線性 → 正規化して次元圧縮(PCA)

1stアプローチ	
線形回帰	ランダムフォレスト
曜日毎に学習	曜日毎に学習
RMSE = 1.880 ウィンドウ期間: 3ヶ月 6次元に圧縮	RMSE = 1.650 ウィンドウ期間: 3ヶ月 ※ローリング回数103回

5. 1stアプローチ

ランダムフォレストと比べ線形回帰は需要も日射量も立ち上がっていく時間帯の誤差が大きい
※この時間帯は需要予測も発電量予測も同様に難しい

ランダムフォレストと比べ線形回帰は日射がない時間帯の誤差が大きい



- ランダムフォレストの精度向上に集中しよう！
- 移動平均値も特徴量に追加してみよう！

日中のうち昼休み時間帯（産業需要が一時的に減少）は両モデルとも精度が良い

ベンチマーク(過去3日間移動平均予測RMSE) = 1.757

6. 2ndアプローチ

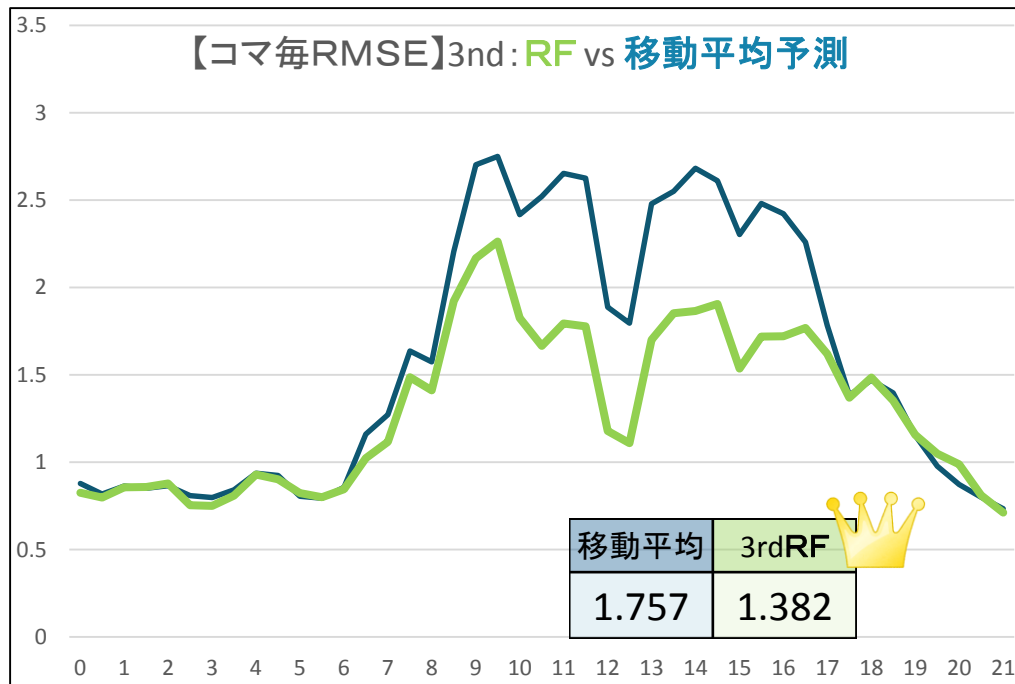
3rdアプローチ	
特徴量	カレンダー情報(月:0～日:6 + 祝:7) → 平日(月～金)と休日(土日祝)に分ける 気象予測値(日射量・気温(差分値)・相対湿度・降水量) × 九州7県分 過去3日間の移動平均値
学習の流れ	コマ毎に学習 ウィンドウ期間過去3ヶ月(90日)をローリング学習
使用モデル	線形回帰、ランダムフォレスト(木の数=150/深さ=5)
処置	気象予測値の多重共線性 → 正規化して次元圧縮(PCA) → 処置無し

1stアプローチ	
線形回帰	ランダムフォレスト
曜日毎に学習	曜日毎に学習
RMSE = 1.880 ウィンドウ期間: 3ヶ月 6次元に圧縮	RMSE = 1.650 ウィンドウ期間: 3ヶ月 ※ローリング回数103回

精度向上

2ndアプローチ
ランダムフォレスト
平日・休日で学習 移動平均値を特徴量に追加
RMSE = 1.382 ウィンドウ期間: 3ヶ月 ※ローリング回数103回

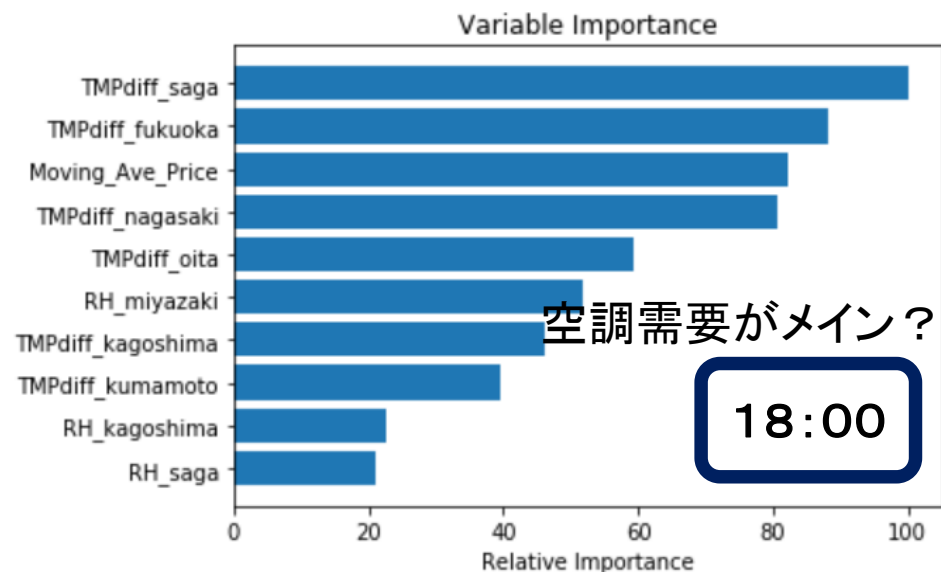
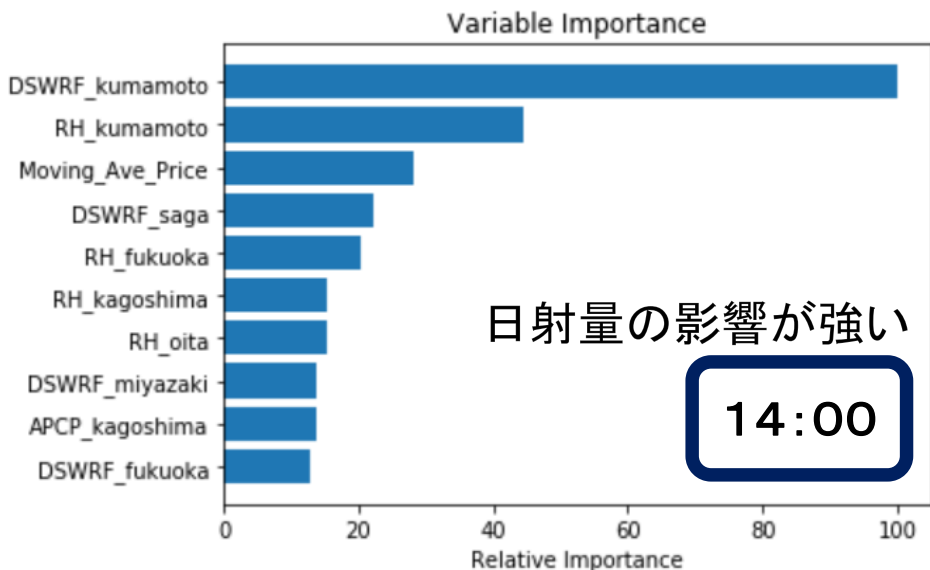
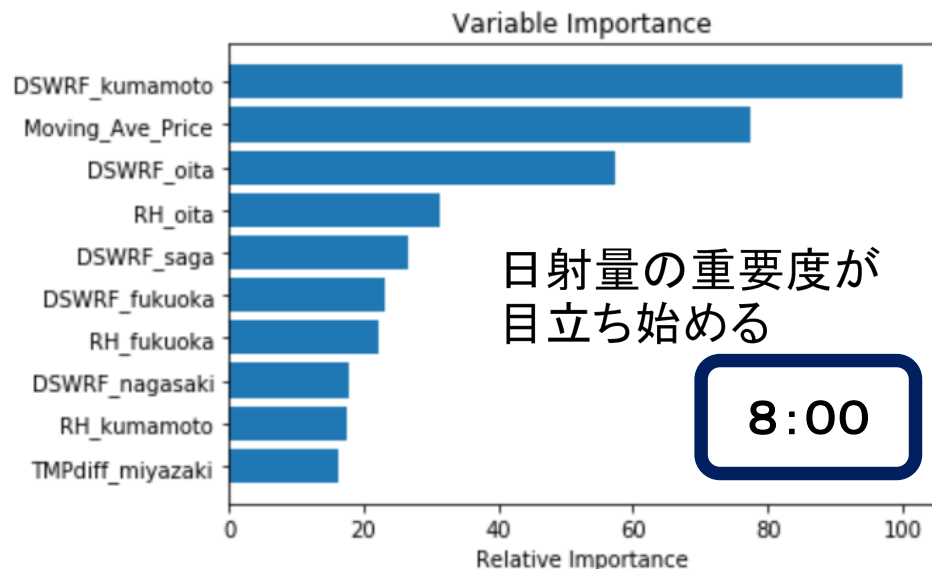
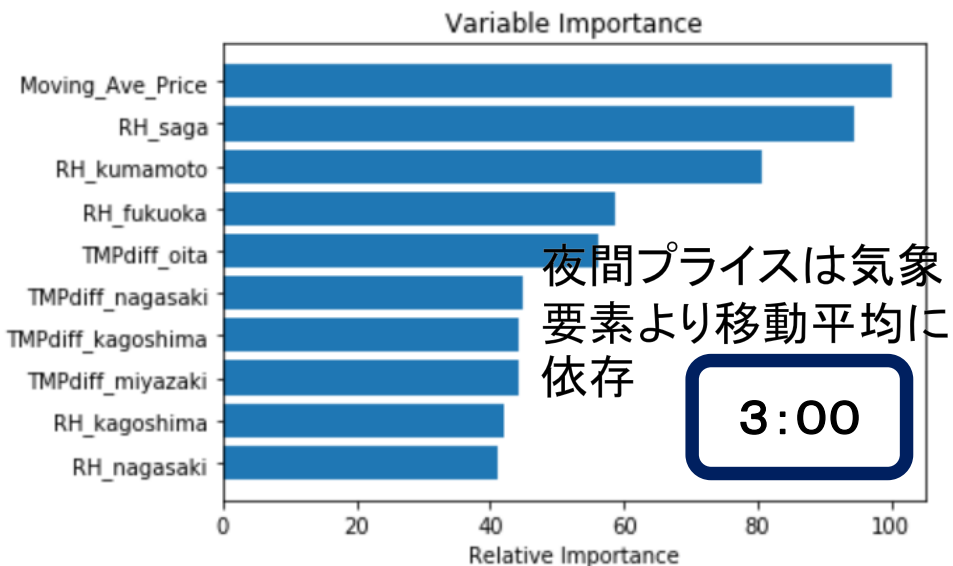
6. 2ndアプローチ



【結論】

- モデルは線形回帰よりもランダムフォレスト
- 曜日毎に特徴量を参照させるより平日・休日の大雑把な分け方が奏効

7. 寄与の高い特徴量『日中価格は太陽光が影響している』

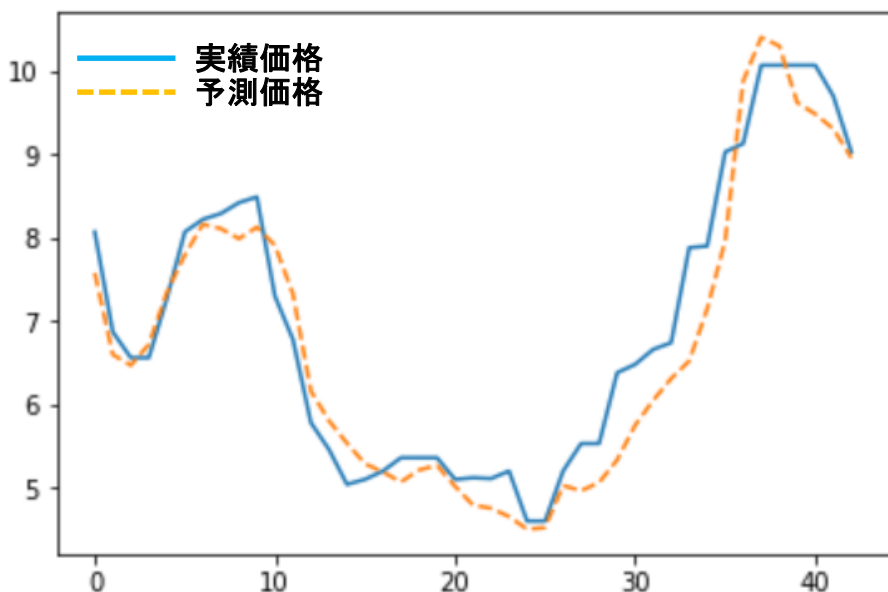


8. うまく予測できた日、ぜんぜんダメだった日

2018年5月12日(土)

	天気(9時)	最高気温	最低気温
福岡	曇	27.1 °C	16.7 °C

66 回目のループRandom Forest RMSE: 0.49202593028452574

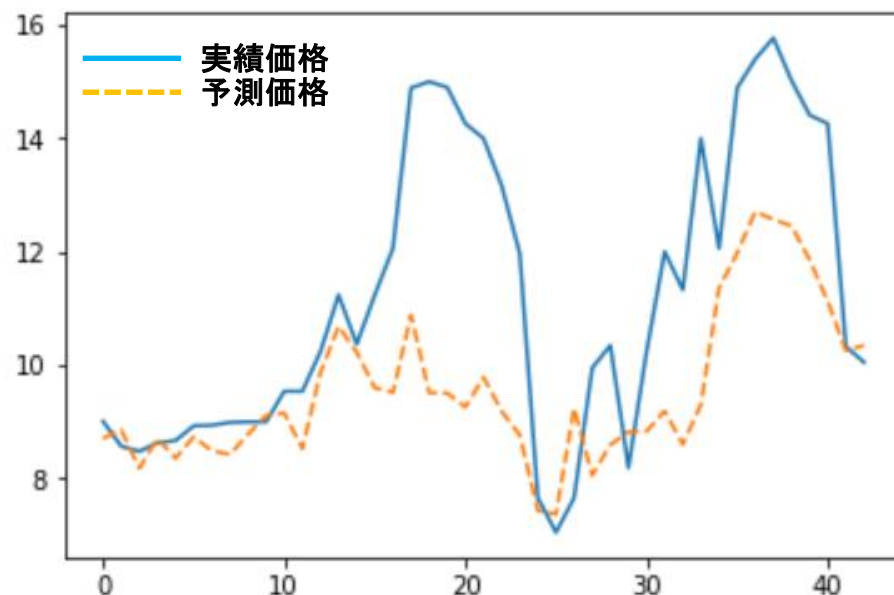


中間期を代表する5月。
天気も安定して価格が暴れず、気象
予測も当たりやすいので、価格予測
の精度も高かった。

2018年3月7日(水)

	天気(9時)	最高気温	最低気温
福岡	曇	15.6 °C	5.9 °C

0 回目のループRandom Forest RMSE: 2.4323429601015465



空調需要が不安定になってくる頃。
気象要因も直近過去の価格でもない、
この頃ならではの反映する特徴量
を見つける必要があるかも。

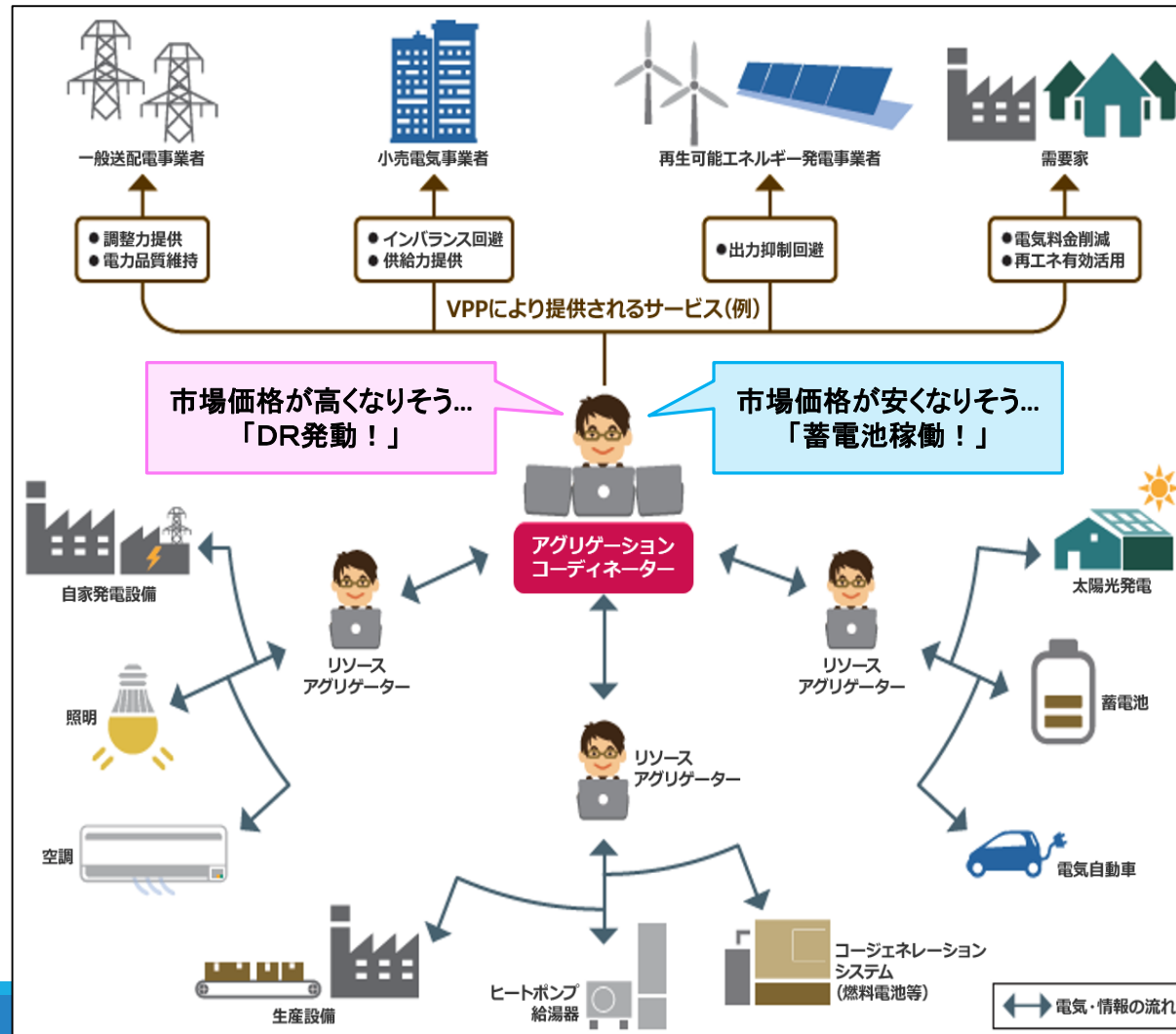
9. 市場価格予測の使用例『VPP運営への活用』

VPP (Virtual Power Plant) とは『仮想発電所』とも言われ、各家庭や工場・オフィス等に設備されている“創エネ”“蓄エネ”“省エネ”のリソースをIT技術によって1つに束ねることで大規模な発電施設のように振る舞わせるシステムです。

例えば①
市場価格が
高くなる場合

節電要請
(DR)

需要を抑える
ことで小売電
気事業者の市
場調達コストを
抑えられる



例えば②
市場価格が
安くなる場合

蓄電池制御

より多く市場調
達して蓄電し
価格が高い時
間帯に放電
(蓄電した電気
を使う)

10. まとめ

- **市場価格で太陽光が大きく影響するとは言えるが、直近過去の価格の影響も無視できない**
 - ☞ 気象予測値だけでなく移動平均も特徴量に加えると精度向上
- **線形回帰よりもランダムフォレストが市場価格予測に良いモデル**
 - ☞ 1年分程度のデータが溜まったらディープラーニング等他の手法も試してみたい
- **曜日毎に特徴量を参照させるより平日・休日のザックリした分け方が奏効**
 - ☞ スパースなデータセットを解消した？
- **日中の気温の影響度が思ったほど大きくなかった**
 - ☞ オペレーションの現場では最高気温を重視していたので、そっちを特徴量にしたほうが精度向上するかも？（特徴量にしたのは時系列気温予測値の20℃差分値）