

人工ニューラルネットワークを用いた短期電力需要予測

－ 国内 10 地域を対象とした検討 －

松尾 雄司*、大谷 公哉*、柴田 智文*、寄田 保夫*、川上 恭章*、永富 悠*

要旨

本稿では旧一般電気事業者の 10 電力供給区域を対象とし、人工知能の一種である人工ニューラルネットワークを用いて翌日の電力需要を予測することを試みた。ここで用いたモデルは日負荷曲線に対して主成分分析を行うとともに選択的アンサンブル平均の手法を用いるものであり、平均誤差率 2.5%弱の良好な予測性能を示している。但し予測性能は季節によって大きく異なり、需要の大きな夏季及び冬季には予測誤差が大きくなる傾向が読み取れる。また予測性能は地域によっても異なり、東北で年平均 1.4%の小さな誤差率を示す一方で、中国や九州では 2.7%と比較的大きな誤差となった。季節による誤差の変化は北海道や沖縄において他地域とは異なる挙動を示し、地域の特性を反映していることが伺える。

今回用いたモデルは各地域の中で 1 地点のみの気象データを用いる簡易なものであり、より大規模なデータを用いることによって更に予測性能の向上を図ることは可能であると考えられる。但し本モデルの延長上で予測性能を向上させたとしても、夏季や冬季でなお比較的大きな誤差を示すことが予想される。異常気象や天気予報の誤差等をも想定した高度な予測手法を検討することは今後の重要な研究課題である。

1. はじめに

近年人工知能 (Artificial Intelligence: AI) の進化が著しく、専門家のみならず広範な人々の関心を惹くに至っている。中でも注目を集めているのは人工ニューラルネットワーク (Artificial Neural Network: ANN) と呼ばれる手法である。この方法の歴史は長く、第二次世界大戦前後から数値計算機の進歩に平行して研究が続けられてきた。ただし深層への学習が難しいことなどから、「冬の時代」と呼ばれる停滞期を数回経たのちに、手法の改良や計算機の能力向上、データの大規模化などとともに大きな成果が開花することとなった。深層への学習を効率的に行う多層の ANN は「ディープラーニング」の手法とも呼ばれ、大量のデータを学習することで人間以上の能力をも発揮し得るものである。2017 年には多層 ANN を用いた「アルファ碁」が囲碁のトップ棋士と対局して全勝し、この手法の有効性を強く印象付けることとなった。本稿ではこの ANN を用いて、日本の電力供給区域ごとの電力需要データをもとに短期の電力需要予測を行う方法について検討する。

短期需要予測 (Short-term Load Forecasting: STL) は電力供給を安定的に行うために重要な課題であり、その手法については古くから研究が行われている。従来、この問題は自己回帰移動平均 (Autoregressive moving average: ARMA) モデルや自己回帰和分移動平均 (Autoregressive integrated moving average: ARIMA) モデルなどの統計的手法が多く用いられてきた。比較的最近では、過去のデータベースをもとに類似日を自動的に特定し、需要予測を行う手法なども公表されている¹⁾。また、TESLA 社では階層化された非線形回帰をベースとし、時系列フィルタリングによって誤差を修正するモデルを開発し、提供しているという²⁾。一方で電力各社は実際の日々

* (一財)日本エネルギー経済研究所

の運用のために電力需要予測を実施しており、ここでは気温、湿度、日射量等の気象データや過去の類似日の需要データ等をもととして、場合によっては大口需要家からの聞き取りをも行い、精度の高い需要予測を行っている³⁾。また近年では太陽光発電設備の導入拡大が著しいことから、エリア内の設備と日射量の正確な分布をもとに、精度の良い太陽光発電量予測を行うことも重要な課題となっている⁴⁾。

これらの取組みが継続的に行われている一方で、最近では機械学習を用いた需要予測が多く試みられるようになってきた。2003年に公表されたレビュー論文⁵⁾では既に人工知能を用いた数十の研究事例が紹介されており、その後十数年の間に更に夥しい数の論文が公表されている。機械学習の手法としてはサポートベクターマシン(SVM)やランダムフォレスト(RF)を用いたものもあるが、特にここ数年の間、少なくとも研究例の数として多いのはANNである。このような予測問題を解くに当り、ANNは少しの工夫で予測性能を大きく向上させ得ることが広く認識されており、課題に合わせてさまざまな工夫がなされている。ここでいう「工夫」とはモデルの入力データを拡充する、パラメータを最適化する、といった基本的な事項とともに、単純な(フィード・フォワード型の)ANNではなく、再帰的ニューラルネットワーク(Recurrent Neural Network: RNN)や畳み込みニューラルネットワーク(Convolutional Neural Network: CNN)、極限学習器(Extreme Learning Machine: ELM)、放射基底関数(Radial Basis Function: RBF)といった応用型のニューラルネットワークを用いることや、遺伝的アルゴリズム(Genetic Algorithm: GA)、粒子群最適化(Particle Swarm Optimization: PSO)、重力探索アルゴリズム(Gravitational Search Algorithm: GSA)、適応ニューロファジー推論システム(Adaptive Neuro-Fuzzy Inference System: ANFIS)といった手法と組み合わせて予測モデルを構築することなどが含まれる¹⁾。例えばRNNについてはその幾つかの変種に対して最適なパラメータを探索し、予測性能の優劣を比較することで、STLFのみを対象とした一冊の書籍が出版されている⁶⁾。

これらの多くの手法の優劣を比較することは容易ではない。その第一の理由として、多くの検討ではある特定のデータセット(地域及び日・時間)のみを対象として評価を行っていることが挙げられる。このため文献に示された誤差指標等の数値を相互に直接比較できないことは当然であるが、そもそも後述するようにモデルの予測性能は対象とするデータセットによって大きく変化するものであり、ある一つのデータセットで良い性能を示す手法が、他のデータセット、例えば同じ地域であったとしても異なる季節の電力需要を必ずしも

精度良く予測できるとは限らない。このため、特定のデータセットを離れて、方法自体としてある一つの手法が他の手法よりも優れていると主張することは容易ではない。第二に、同じ手法に基づいたモデルであっても、そのパラメータ、例えば後述する層やニューロンの数、もしくは活性化関数の形状、正則化項の有無といった細かな差異によって予測性能は変化する。原理的には最も予測性能が良くなるパラメータ等を選ぶべきであるとは言えるが、上述の通り予測性能の優劣自体が必ずしも一意的には定まらない上に、全てのパラメータを最適化することはかなりの時間を要する作業であり、多くの場合は試行錯誤によってそれらが決定されている。このため二つの手法を比較する際に、比較者が推奨したい手法のみがその課題により良く適合するように調整され、それと同等の労力が他方のモデルに対して払われていない可能性が常に存在する。

このような状況の中で、同一の条件の下で多数の参加者が予測を行い、その優劣を比較する試みも行われている。例えば米ノースカロライナ大学シャーロット校のTao Hong博士らが主催するGlobal Energy Forecasting Competition (GEFCom)^{7,8)}は2012年、2014年及び2017年に開催され、確率論的/決定論的/階層的な電力需要や太陽光・風力発電量などの予測を対象として評価を行っている。またフランスの送電事業者Réseau de transport d'électricité (RTE)⁹⁾と日本の東京電力¹⁰⁾はそれぞれ翌日の電力需要予測を対象としたコンテストを実施している。これらの試みもある特定のデータセットを対象としている以上、真に客観的な比較を可能にすると

¹⁾ ここで挙げたものはここ2、3年の間に公表された文献に見られる検討事例の「一部」であることに留意されたい。

は言い難いものの、これまで困難であったモデル間の相互比較を行う試みもある程度は進んできているとは言える。

本稿では ANN による短期需要予測モデルを作成し、それを用いて国内の 10 電力供給区域ごとに予測性能を評価した。後述する通りここで用いたモデルはそれなりに高い予測性能を有すると思われるものの、過去の電力需要値と暦データの他に、域内の 1 地点のみの気象データ（天気概況及び最高・最低気温）を予測に用いる簡素なモデルであり²、これが「最良」のものであると主張するものでは決してない。ここでは手法の改良やデータの拡充による更なる予測性能向上の可能性を念頭に置きつつ、比較的素直で簡素な工夫を行ったモデルによって電力供給区域ごと・季節ごとの予測性能を評価し、ANN による予測の特徴を把握することを目的としており、これを踏まえた今後の更なる研究の進展を期待するものである。

2. 人工ニューラルネットワーク (ANN) を用いた短期電力需要予測モデル

ANN は入力のベクトルと目標値（スカラーまたはベクトル）との間の非線形な関係を、多量のデータを用いて機械に学習させる手法の一つである。その概念図は図 2-1 のように表される。

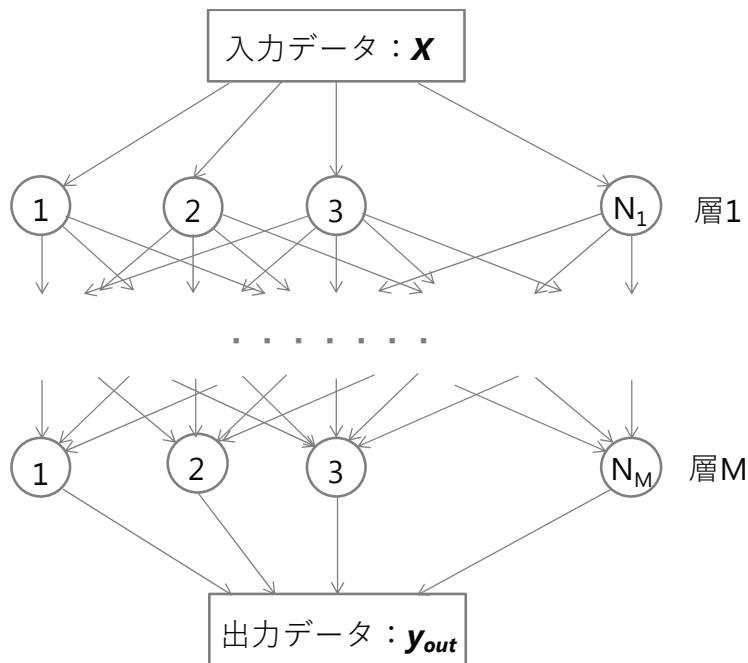


図 2-1 多層人工ニューラルネットワークの概念図

ここでは、入力層と出力層の間に M 層の中間層（隠れ層）が設定され、 n 番目の層は N_n 個のノード（ニューロンと呼ばれる）によって構成される。即ち M や N_n の大小によって、モデルの規模（複雑さ）が決定されることになる。このモデルに対して入力データ X と出力データ Y との多数の組を与えてモデルを「学習」させた上で、更に新たな入力データ X_F を与えることにより、未知の値 Y_F を推定しようとする試みが ANN による将来予測である。

² これに対し、例えば東京電力・「第 1 回電力需要予測コンテスト」で最優秀賞を受賞した（株）東芝のモデルは東京電力管内の多地点の AMeDAS データを用いており、そこからスパースモデリング技術によってデータを抽出し、利用するものであるという。

https://www.toshiba.co.jp/about/press/2017_11/pr_j0801.htm

ANN の特徴は大量のデータを用いて学習を行うことにより、幅広い変化の可能性を捉えた精緻なモデル表現が可能となることである。一般的には学習に用いるデータは多ければ多いほど良いが、実際に過去の実績値をもとに将来推計を行う場合には、学習データの数が制限される場合が多い。例えば今回試みる電力需要予測では、仮に 10 年間の需要データを集めたとしてもその数は 4,000 日に満たず、機械学習の観点からは決して多くのデータが利用可能であるとは言えない。このような場合、限られたデータからいかに効率的に学習を行うかが予測の性能を左右することになる。

また、学習が行われた後のモデルは基本的にはブラックボックスであるため、その中身を人間が理解することは一般的には難しい。このため、ANN はあくまでも実用上の道具としてのみ用いることができるものであり、それによって事象の理解を深めることは一般的には難しい、ということには注意する必要がある。

今回の試算では電力会社の実際の運用を想定し、 d 日の朝、具体的には午前 8 時以前に $d+1$ 日の電力需要の 24 時間値を予測することを試みた。即ち利用可能なモデルへの入力データは全ての日の暦データ（年、月、日、曜日及び祝日）の他に、 d 日の朝 7 時までの電力需要実績値と、 $d-1$ 日までの気象データ（天気概況及び最高・最低気温）実測値、並びに d 日朝時点で得られる翌日までの気象予報データである。気象データとしては一般的には区域内の多地点のデータが利用可能であるが、ここでは旧一般電気事業者の本社・本店の所在地 1 地点のみのものを利用した。

気象予報データ、例えば気温の予報値は一般的には 24 時間値が利用可能である。また電力需要には湿度が大きく影響することも広く知られており¹¹⁾、これらを用いることで有意に予測性能を向上させ得ることが期待できる。しかし今回、過去のデータを用いたモデルの検証を行うに当たって最高気温・最低気温の予報値のみが利用可能であったことから、今回はこれを利用して予測を行うモデルを作成し、検討を行った。

予測モデルの出力値は $d+1$ 日の 24 時間値であるが、本稿では主成分分析を用いて日負荷曲線を分析するとともに、選択的アンサンブル平均をとることで予測誤差を低減させるモデルを作成・利用した。これにより、本稿で示す結果は単純な ANN を用いた場合に比べて、有意に良好な予測性能を示している。モデルのより詳細な説明については附録 1 を参照されたい。

3. 使用データの概要及び試算条件

3-1 使用データ

3-1-1 電力需要データ

本分析では ANN を用いて旧一般電気事業者各社管内 10 地域における電力需要の予測を行った。分析にあたっては各社が公表した電力需要データを用いている。用いたデータの期間等を表 3-1 に示す。基本的には 2012 年 1 月 1 日から 2018 年 6 月 30 日までのデータを用いてモデルの学習・検証・評価等を行っているが、入手可能であったデータの制約から、沖縄のみ 2016 年 4 月 1 日以降のデータを用いて推計した。

各地域の電力需要は地域特有の経済活動や気象条件などを反映しており、附録 2 にまとめるとおり、季節ごとの変化、日単位の変化について各地域それぞれの特色がある。

表 3-1 電力需要データ一覧

地域	収録期間	データ 日数	データ 欠損日数	欠損日
北海道	2012/1/1-2018/6/30	2,341	32	2015/1/24, 2015/2/28, 2015/3/1, 2015/8/8 ~11, 2016/2/5~2016/2/29
東北	2012/1/1-2018/6/30	2,341	32	2015/1/24, 2015/2/28, 2015/3/1, 2015/6/15(一部), 2015/11/30, 2016/2/5~ 2016/2/29, 2016/3/14(一部), 2016/3/31
東京	2012/1/1-2018/6/30	2,344	29	2015/1/24, 2015/2/28, 2015/3/1, 2016/2/5 ~2016/2/29, 2016/3/31
中部	2012/1/1-2018/6/30	2,336	37	2015/1/24, 2015/2/28, 2015/3/1, 2015/4/30(一部), 2015/8/27(一部), 2015/12/5(一部), 2016/2/5~2016/2/29, 2016/3/3(一部), 2016/3/6(一部), 2016/3/23(一部)
北陸	2012/1/1-2018/6/30	2,345	28	2015/1/1~1/4, 2015/1/10~1/12, 2015/1/24, 1/25, 2015/2/1, 2015/2/7, 2/8, 2015/2/11, 2015/2/14, 2/15, 2015/2/21, 2/22, 2015/2/28, 2015/3/1, 2015/3/7, 3/8, 2015/3/14, 3/15, 2015/3/21, 3/22, 2015/3/28, 3/29+
関西	2012/1/1-2018/6/30	2,347	26	2015/2/28, 2016/2/5~2016/2/29
中国	2012/1/1-2018/6/30	2,303	70	2015/1/1~1/4, 2015/1/10~1/12, 2015/1/17, 1/18, 2015/1/24, 1/25, 2015/1/31, 2015/2/1, 2015/2/7, 2/8, 2015/2/11, 2015/2/14, 2/15, 2015/2/21, 2/22, 2015/2/28, 2015/3/1, 2015/3/7, 3/8, 2015/3/14, 3/15, 2015/3/21, 3/22, 2015/3/28, 3/29, 2015/4/1~4/14, 2015/10/20, 2016/2/5~2016/2/29
四国	2012/1/1-2018/6/30	2,372	1	2012/1/1
九州	2012/1/1-2018/6/30	2,368	0	
沖縄	2016/4/1-2018/6/30	820	1	2016/4/1

(出所) 各社HP、公表資料より作成。データの欠損は取得漏れによる。

3-1-2 気象データ

気象データとしては、旧一般電気事業者の当社・本店の所在となる各都市(表3-2)につき、気象庁が公表する実績データ及び予報データを用いた。ここでは天候を示す変数(その日の天気概況が「晴」「曇」「雨」「雪」

の各文字を含むか否か) 及び最高気温・最低気温の値を用いている。

表 3-2 気象データ対象都市

地域	都市	地域	都市
北海道	札幌	関西	大阪
東北	仙台	中国	広島
東京	東京	四国	高松
中部	名古屋	九州	福岡
北陸	富山	沖縄	那覇

3-2 試算条件及び評価指標

3-2-1 試算条件

本稿では表 3-1 に示す電力需要データを用いて、2017 年 7 月から 2018 年 6 月までの期間を対象とした予測を行い、その予測値を実績値と比較して予測誤差を評価した。

人工ニューラルネットワーク等によって予測を行う際には、その時点で得られているデータを学習データと検証データに分けて評価を行うことが一般的である。学習データは文字通りモデルを学習させるために用いられる一方で、学習したモデルを検証し、ハイパーパラメータの設定などを行うために検証データが用いられる。本稿で示す手法では、選択的アンサンブル平均をとる際に検証データを用いている。

ここでは予測期間直前の 1 ヶ月間を検証期間とすることとした。例えば 2018 年 2 月の予測を行って予測誤差を評価する場合、2012 年 1 月 (沖縄の場合は 2016 年 4 月) から 2017 年 12 月までのデータを学習データ、2018 年 1 月のデータを検証データとしてモデルの学習を行う。このようにして得られたモデルに対し、例えば 1 月 31 日朝時点で得られるデータを入力して 2 月 1 日の電力需要を予測し、2 月 1 日朝時点で得られるデータを入力して 2 月 2 日の電力需要を予測し、・・・と繰り返すことにより 2 月 28 日までの需要を予測して、それらの予測値を実際の需要値と比較した。

なお同じ構造の ANN 及び全く同じ学習データであっても、学習前の初期値 (一般的には乱数で与えられる) によって異なる学習をするため、それによる予測値 (従って予測性能) も異なるものとなる。このため本稿では初期値を変えて同一の計算を 15 回繰り返し、その予測誤差の平均値によって評価を行った。

3-2-2 予測性能の評価指標

予測性能を示す指標として、次に示す平均自乗誤差 (Root Mean Square Error: RMSE) 及び平均絶対誤差率 (Mean Absolute Percentage Error: MAPE) を採用した。

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_t (f_t - y_t)^2} \quad (1)$$

$$\text{MAPE} = \frac{1}{n} \sum_t \left| \frac{f_t - y_t}{y_t} \right| \quad (2)$$

ここで f_t は予測値、 y_t は実績値であり、 n は予測期間におけるデータの個数である。RMSE は予測対象データ、今の場合各地域内の電力需要と同じ単位 (万 kW) をもつ誤差指標であり、各社別の電力需要の中で誤差の規模感を見るために有用である。一方で MAPE は誤差率 (%) を示しており、異なる地域間での予測性能を比べる

ために有用である。なお式(1)及び(2)からわかる通り、RMSE は MAPE に比べて大きな誤差に対してより大きなペナルティを与えるため、両者による予測性能の順位付けは常に一致するとは限らない。

4. 試算結果及び考察

4-1 試算条件の変化による影響

4-1-1 予測結果例

予測結果の例を図 4-1～図 4-4 に示す。ここでは東京電力管内を対象として、夏季（2017 年 8 月 21 日～30 日）、秋季（2017 年 9 月 1 日～10 日）、冬季（2018 年 2 月 15 日～24 日）、春季（2018 年 4 月 1 日～10 日）の翌日電力需要を予測した結果を実績値とともに示している。図中には予測結果が一本の曲線で描かれているが、実際には上述の通り当日の朝までに得られている情報のみをもとに翌日の 24 時間値を予測し、それをつなぎ合せて示している。またこれらの図は 15 回試みた結果の中での代表的な 1 例のみを示したものであることに留意されたい。

ここに示される通り同じ東京電力管内であっても予測誤差は季節によって大きく異っており、夏季及び冬季には RMSE がそれぞれ 195 万 kW 及び 174 万 kW と大きく、秋季及び春季にはそれぞれ 79 万 kW 及び 60 万 kW と小さい。例えば夏季の 2017 年 8 月 28 日には予測値は実績値を大きく下回っているが、この日の前日に公表された予想最高気温は 28°C、天気予報は「曇」となっており、実際の 31.4°C、「曇後晴」と大きく異っていた。また 8 月 30 日には予想最高気温は 33°C でほぼ実際の値と同じであったものの、予想最低気温が 27°C と、実際の 24.1°C を大きく上回っている。また天気予報を外れて雨が降ったこともこの日の特徴である。

冬季で大きく予測が外れている 2018 年 2 月 22 日は予報を外して雪が降った日であり、最高気温も 5.7°C と、予想の 9°C を大きく下回っている。次の 23 日も天候は「みぞれ一時曇後晴」となっており、最高気温 8.1°C は予報値 11°C を大きく下回っていた。このように、天気予報の誤差による予測誤差への影響は大きい。一般的には夏季には実際の気温が予報値よりも高い、もしくは予報に反して晴れた場合に電力需要は予測値を上回り、冬季には逆に実際の気温が予報値よりも低い、もしくは予報に反して雪が降った場合などに電力需要が予測値を上回る。

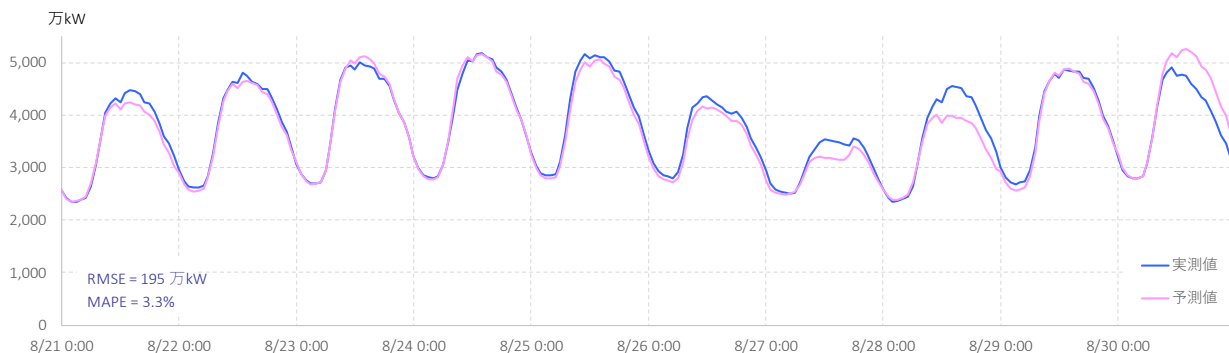


図 4-1 東京電力管内の需要予測例（2017 年 8 月 21 日～30 日）

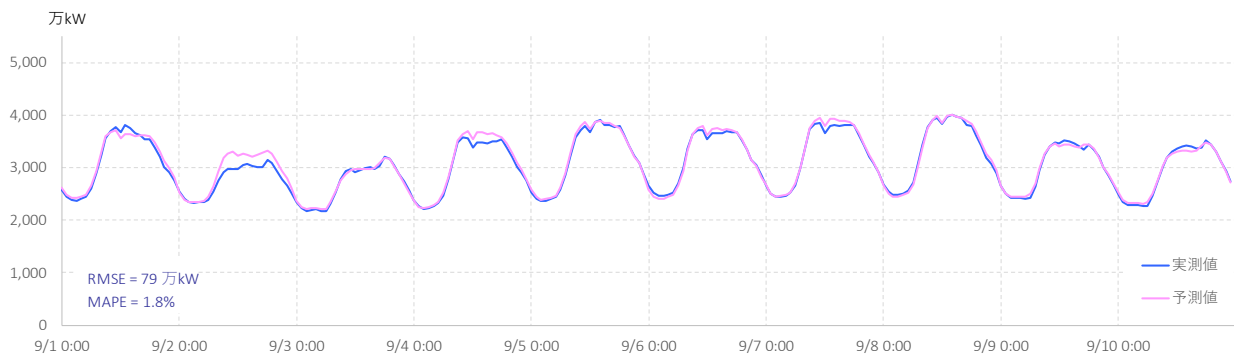


図 4-2 東京電力管内の需要予測例 (2017 年 9 月 1 日～10 日)

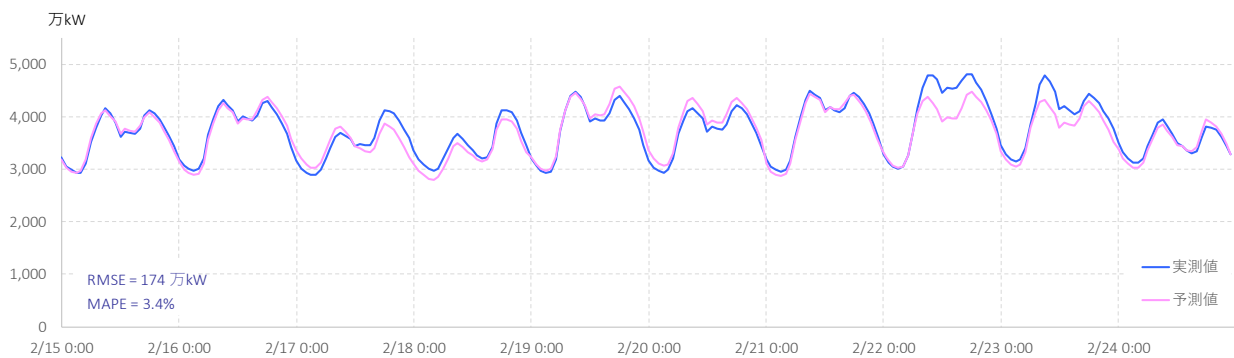


図 4-3 東京電力管内の需要予測例 (2018 年 2 月 15 日～24 日)

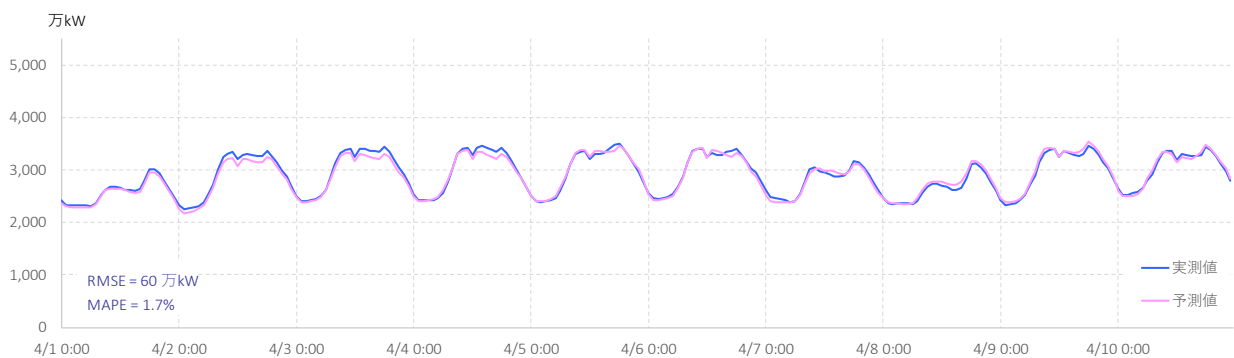
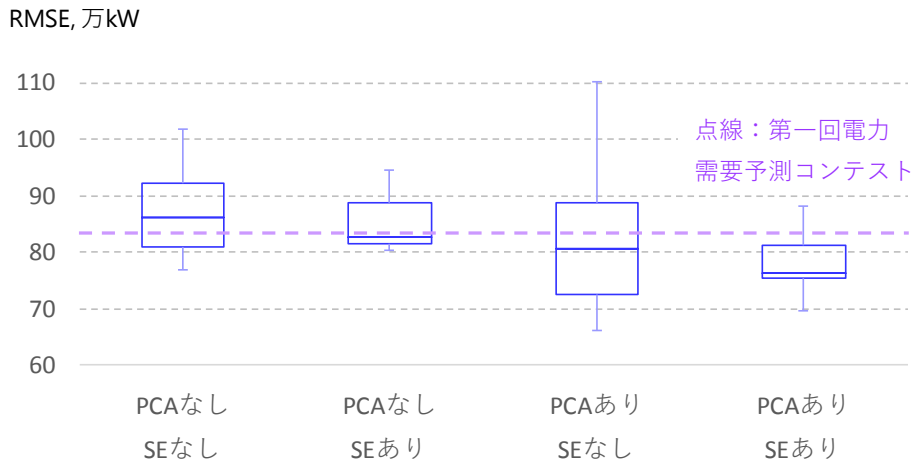


図 4-4 東京電力管内の需要予測例 (2018 年 4 月 1 日～10 日)

ここで、図 4-2 に示す 2017 年 9 月 1 日～10 日のうち、最初の日を除く 2 日から 10 日は東京電力による「第 1 回電力需要予測コンテスト」の予測対象期間と一致しており、図に示す例ではこの 1 日を除いた期間の RMSE は 77.8 万 kW となっている。この期間を対象とした RMSE をそれぞれ 15 回の試行について箱髭図で示すと図 4-5 の通りとなる。主成分分析を行わず、選択的アンサンブル平均も行わない素朴な方法では、15 回の RMSE の中央値は 86 万 kW、平均値は 87 万 kW となる。それに対し、選択的アンサンブル平均をとった場合、及び主成分分析を行った場合には中央値はそれぞれ 83 万 kW 及び 81 万 kW、平均値はそれぞれ 85 万 kW 及び 83 万 kW となる。両方行った場合には誤差は更に低減し、中央値 76 万 kW、平均値 78 万 kW となる。第一回電力需要コンテストにおいて最優秀賞を受賞した (株) 東芝による予測誤差が 83.49 万 kW であったことを考慮すると、少

なくともこの予測期間に対しては、本手法によって良好な予測が可能であったと考えられる。但し予測誤差は季節によって大きく変化するものであり、後述の通り9日間の小さなデータセット上のみで行われた比較は、必ずしもモデルの予測性能自体を適切に評価しているとは限らない。つまり本手法が常に他の方法に比べて良好な結果を示すとは限らないことには、十分に注意する必要がある。



PCA：主成分分析、SE：選択的アンサンブル平均

図 4-5 東京電力管内の需要予測誤差 (RMSE：2017年9月2日～10日)

4-1-2 計算条件による予測誤差の変化

計算条件の差による予測誤差の変化を図4-6 (RMSE) 及び図4-7 (MAPE) に示す。ここでは以下の5つの条件のもとに計算を行っている。

- S：2016年4月1日からのデータで学習、1地点（東京）の気象予報データを用いて予測
- M：2012年1月1日からのデータで学習、1地点（東京）の気象予報データを用いて予測
- M2：2012年1月1日からのデータで学習、2地点（東京・宇都宮）の気象予報データを用いて予測
- MT：2012年1月1日からのデータで学習、1地点（東京）の気象データ実績値を用いて予測
- L：2008年1月1日からのデータで学習、1地点（東京）の気象予報データを用いて予測

ここでは上述の通り一つの月について15の異なる初期条件のもとに試算を行い、その予測誤差の平均値のみを示しているが、実際には初期条件に応じて各月RMSEについては3～4万kW程度、MAPEについては0.1%程度の標準誤差で値がばらついていることに留意されたい。

S、M及びLの差は学習に用いるデータセットの大きさの差であり、一般的にはより大きなデータを用いて学習を行った方が予測性能が向上すると考えられる。但し遠い過去と直近とで電力需要の変化が異なる挙動を示す場合には、過去のデータを学習に含むことで予測性能が悪化することも当然あり得る。例えばLは2011年3月の東日本大震災以前のデータをも含んでおり、実際に震災前後で国民の電力消費行動には大きな変化があったこと、また特に2011年には他年と異なる節電行動があったことなどを考えると、2011年以前のデータが現在の電力需要の変化とどの程度関係しているかは、アプリアリには明確でない。

図4-6と図4-7とにおいてSとMを比較すると、全ての場合についてSよりもMの方が誤差が小さい。年間平均でみるとSのRMSE 133万kW、MAPE 2.7%に対し、MではRMSE 116万kW、MAPE 2.3%とかなり誤

差が小さくなっている。このことから、Sのように2年間程度のデータを用いて学習させるだけでは必ずしも十分ではなく、より長期のデータを使うことにより予測精度を向上させ得ることがわかる。

一方でMとLとを比較すると、例えば1月や9月にはLの方が誤差が小さくなっている一方で、2月や12月には逆にMの方が誤差が小さくなっている。通年平均でみるとRMSE及びMAPEは概ね同程度であり、統計的に有意な差が見られない。このように、2011年以前のデータを学習させることによって予測性能は有意に向上しない。但しこれが上述のような震災前後の電力消費行動の変化によるものであるか、或いは単に6年間程度以上の長期のデータを含めても予測性能は向上しない、ということの意味するのかについては、更に検討が必要である。

MとM2の差は多地点化の影響を示す。一般的には、より多くの地点のデータを用いて学習を行うことにより予測誤差が小さくなると想定される。但し図4-6及び図4-7を見ると多くの月でM2の方が誤差が小さくなっている一方で、1月～3月や9月のように逆に誤差が大きくなる月もある。通年平均ではRMSEは116万kWから115万kWへと僅かに低下するが、この差についても統計的に有意であることは、少なくともこの結果のみからは示されない。

MTは気象予測データの代わりにその日の実測の気象データを用いて計算を行った場合である(従ってこれは「予測」ではない)。ここでは当然ながら誤差が大幅に低下しており、年間平均のRMSEはMの116万kWに対して90万kW、MAPEは2.3%に対して1.9%となっている。但しこのケースにおいても夏・冬の方が春・秋よりも誤差が大きくなる傾向は変わらない。即ち、天気予報の誤差は電力需要予測の誤差の大きな要因ではあるものの、それ以外の何らかの理由によっても夏・冬の予測誤差が大きくなっていることが推測される。

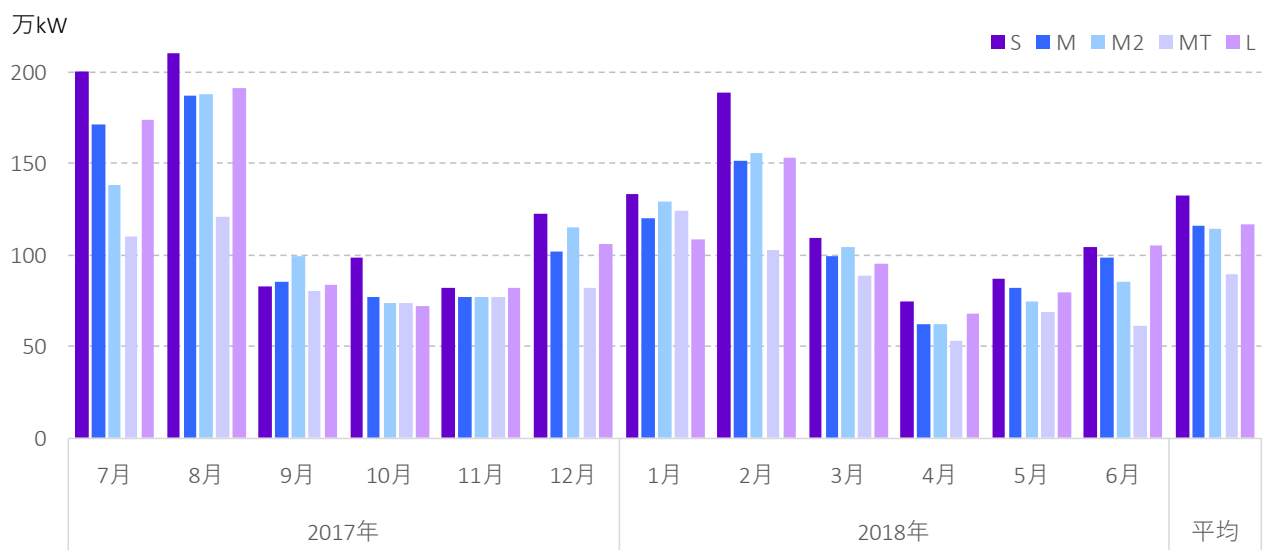


図 4-6 計算条件による予測誤差の変化 (RMSE:東京電力管内)

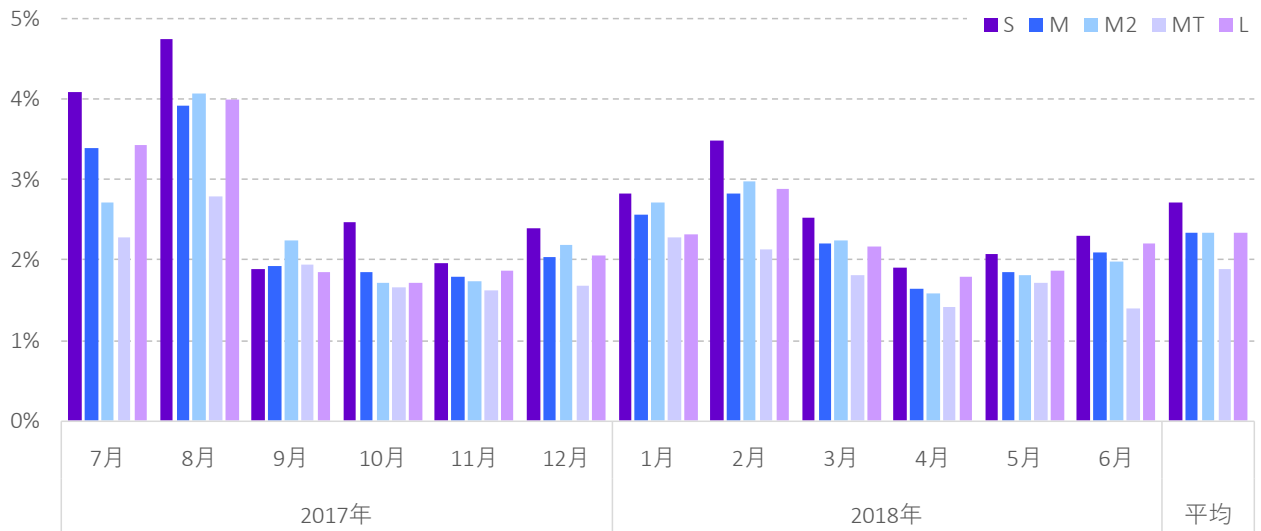


図 4-7 計算条件による予測誤差の変化 (MAPE：東京電力管内)

なお図 4-6 及び図 4-7 にも示されるように、二つのモデル、例えば M と M2 との予測誤差の優劣は月によって大きく変化することには十分に注意する必要がある。例えば 2018 年 1 月を対象として評価した場合、M よりも M2 の方が誤差が大きくなる。これはランダムな初期値の与え方によって「たまたま」そうなったのではなく、多数回の計算を行うと統計的にも有意にそのような結果となる。しかしそれは M の方が M2 よりも「良い」モデルであることを意味するわけではない。即ち、例えば 7 月には M よりも M2 の方が逆に有意に誤差が小さくなり、通年平均では上述の通り有意な差は見られない。このことから、電力需要予測において異なる手法やモデルの優劣を評価する場合、小さなデータセットの上でのみ行った評価は妥当でない可能性が高いと考えるべきであろう。即ちモデルの相互比較を行う際には、可能な限り大きなデータセットの上で評価を行うことが望ましい。

一方で M と M2 の比較に限って言えば、M2 の誤差が大きくなることがあるのは宇都宮の気象を入力データとすることが必要以上に擾乱を与えるためと考えられることから、例えば更に多地点の気象条件を用いる、或いは人口や電力需要などで重み付けを行う等の処理をすることにより、年間を通じて有意に予測性能を高め得る可能性もあると思われる。

4-2 地域ごとの予測誤差

4-2-1 予測結果例

本節では例として、2018 年 6 月 20 日～29 日の電力供給区域（10 地域）ごとの予測結果を示す。ここでは前述の通り沖縄以外の全ての地域について 2012 年からのデータを用いて学習を行っており、また各地域 1 地点のみの気象データをもとに予測を行っている。

(1) 北海道

北海道電力管内の 6 月 20 日～29 日の平均電力需要は 307 万 kW、最大は 29 (金) 午後 4 時の 362 万 kW、最小は 26 (火) 0 時の 258 万 kW であり、この期間は約 100 万 kW の比較的狭いレンジで電力需要が変動していたため予測がし易かったといえる。そうした背景もあり、平均 RMSE=7 万 kW と良好な予測結果となった。一方で、予想最高気温 15°C に対して最高気温が 26.2°C (誤差 11.2°C) であった 23 (土) は、RMSE=9.5 万 kW、MAPE=2.4% と他の日に比べ精度が悪化した。特に気温の予測が大きく外れたと考えられる 10 時～17 時は RMSE=15 万 kW と長時間にわたって過小な予測となっている。また、短時間で最も予測が外れたのは 26 (火)

11時で、最大電力需要345万kWに対し約23万kW過小な予測をした。これは午前中に突如発生した電力ピークを予測することができなかったことによる。同様の予測外しは20(水)、22(金)、27(水)にも確認された。いずれも1日の予想最高・最低気温に大きな誤差は確認されておらず、もしこの電力需要ピークが気温に起因する場合は、1時間毎の気温予報データを学習させなければ改善できない可能性がある。

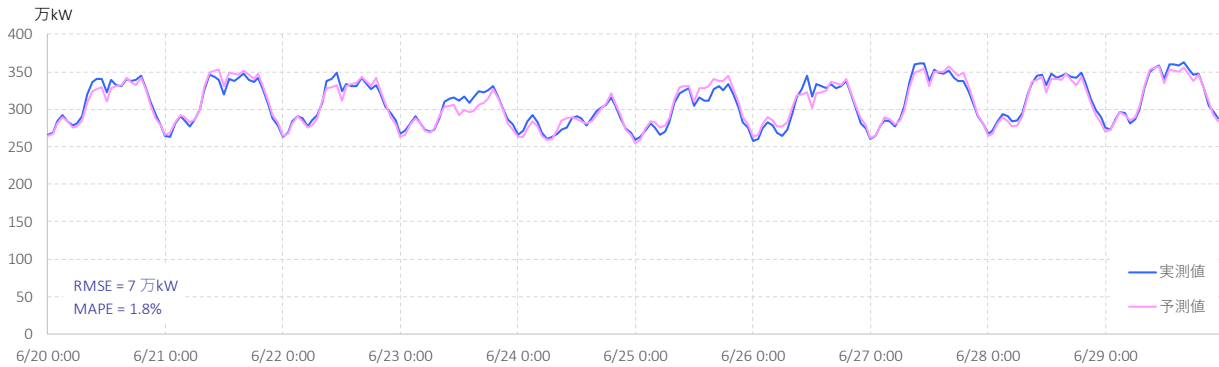


図 4-8 北海道電力管内の予測結果例

(2) 東北

東北電力管内の6月20日～29日の平均電力需要は682万kW、最大は29(金)午前11時の810万kW、最小は24(火)1時の594万kWであり、この期間は約200万kWの比較的狭い範囲で電力需要が変動していたため予測がし易かったといえる。そうした背景もあって平均RMSE=13万kW、MAPE=1.3%と良好な予測をし、この平均のMAPEはこの期間における予測では10地域中で最高の精度であった。一方で、26(火)、29(金)にはRMSE=18万kW及び25万kW、MAPE=2.1%及び2.7%と実績に比べて過小の予測をした。この2日に共通することは、学習に用いた仙台の気温変化に対して、他の地点、例えば新潟の気温変化が異なる動きをしたという点である。例えば、26は仙台の最高気温は対前日4°C低下であったが、新潟は6.8°C上昇していた。また、29日は仙台の最高気温は対前日4°C上昇したのに対し、新潟は7.6°C上昇していた。いずれもこの過小予測は他の地域での需要増加を予測できなかったことに起因すると考えられ、この様に供給範囲が広く電力需要が点在している東北電力管内では、気象データ1地点では他の地域での需要変化に対応仕切れない可能性が伺える。

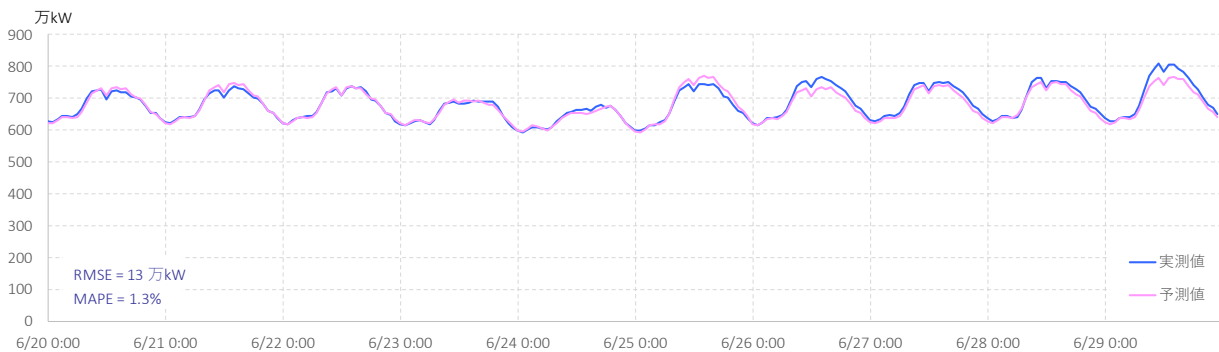


図 4-9 東北電力管内の予測結果例

(3) 東京

東京電力管内の6月20日～29日の平均電力需要は3,316万kW、最大は29(金)午後2時の4,727万kW、最小は24(日)午前5時の2,213万kWであった。モデルはRMSE=81万kWと良好な予測をしているが、27(水)には最大電力需要4,540万kWを4,257万kWと過小に予測している。この日の最高気温31.7°C・最低気温24.9°Cに対して予想最高気温は30°C、最低気温は23°Cと、天気予報の誤差に応じて予測誤差が生じたことが伺える。

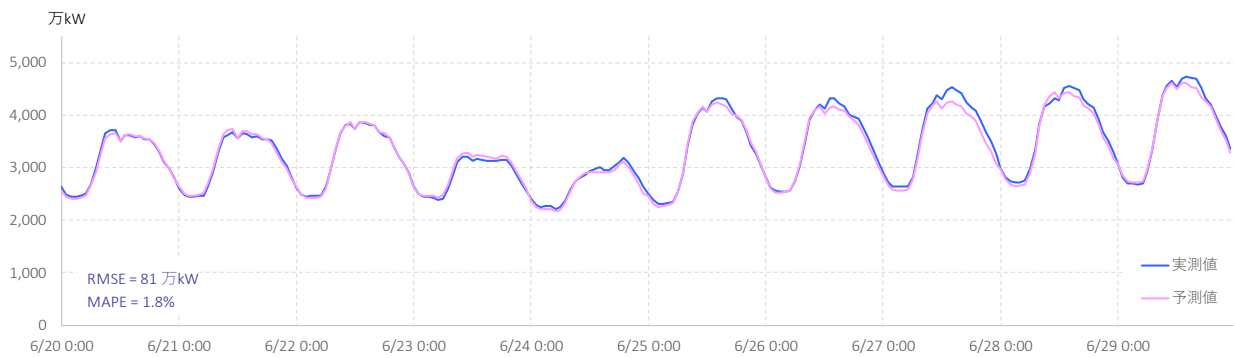


図 4-10 東京電力管内の予測結果例

(4) 中部

中部電力管内の6月20日～29日の平均電力需要は1,578万kW、最大は28(木)午後2時の2,130万kW、最小は24(日)午前6時の997万kWであった。RMSE=45万kWと概ね良好な予測をしているが、25(月)～27(水)の午後においては、ほぼ一貫して需要を過少評価する傾向が見られた(平均67万kW、最大124万kWの過少評価)。当該3日間における最高気温の実績－予報差は平均+1.5°C程度であり、27日の最高気温で+2.5°Cと比較的高い予報差を記録した点を除いては、他期間と比較して顕著な予報差は生じていない。

他方で、最高気温実績に着目すると、23(土)以前の期間は、概ね30.0°C以下で推移したのに対し、24(日)には31.1°C、翌25(月)には33.9°Cと2日連続で当年度最高を更新している。このことから、急激な気温上昇に起因する冷房需要の増加局面において、モデルによる実績－予報差が拡大する傾向が示唆される。

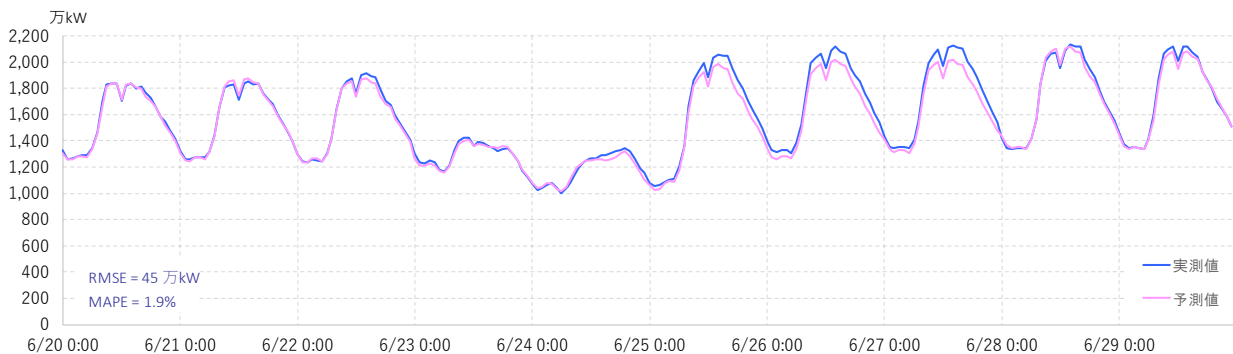


図 4-11 中部電力管内の予測結果例

(5) 北陸

北陸電力管内の6月20日～29日の平均電力需要は333万kW、最大は29(金)午前11時の440万kW、最小は24(日)午前7時の232万kWであった。RMSE=12万kWと概ね良好な予測をしているが、29(金)午後より予想値に過大評価の傾向が見られ、同日午後4時に、実績値401万kWに対し予測値451万kWと期間最大の過大評価が生じている。この日の最高気温30.7°Cに対して予想最高気温は34°Cと、天気予報の誤差に応じて予測誤差が生じたことが伺える。

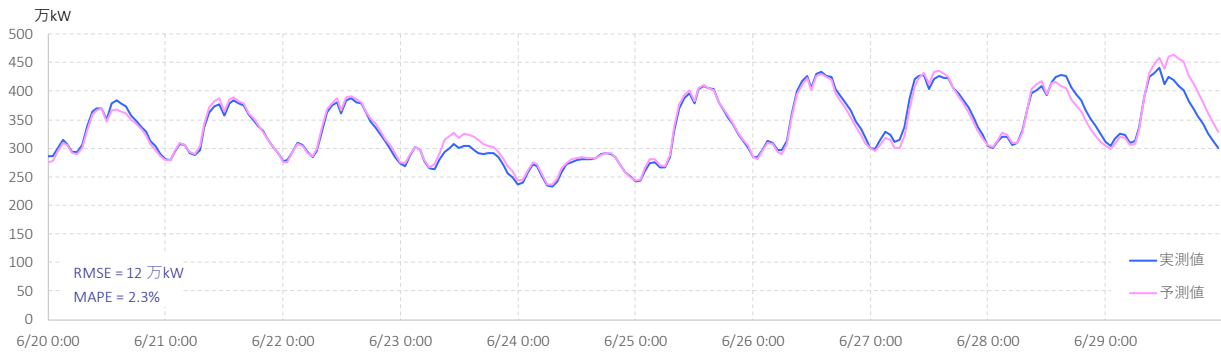


図 4-12 北陸電力管内の予測結果例

(6) 関西

関西電力管内の6月20日～29日の平均電力需要は1,681万kW、最大は28(木)午後2時の2,319万kW、最小は24(日)午前1時の1,160万kWであった。RMSE=70万kW、MAPE=3.0%と予測精度は他地点よりやや劣後しており、特に、25(月)～27日(水)の日中付近(午前9時～午後6時までの間)については、予測値が需要を過少評価する傾向が顕著に見られた。

この要因として、当該3日間における最高気温の実績－予報差の平均が+2.1°C程度と、他の期間の予想差との比較でやや高めに推移したことが考えられるほか、6月24(日)までは概ね30.0°C以下で推移していた最高気温が、25(月)に33.1°Cと当年度最高気温を更新、翌26(火)に31.6°C、続く27(水)に33.5°Cと、急激に上昇にしたことで、冷房需要が急速に増加、それに伴う急激な需要の変化を、モデルが正確に捕捉できなかったことが考えられる。

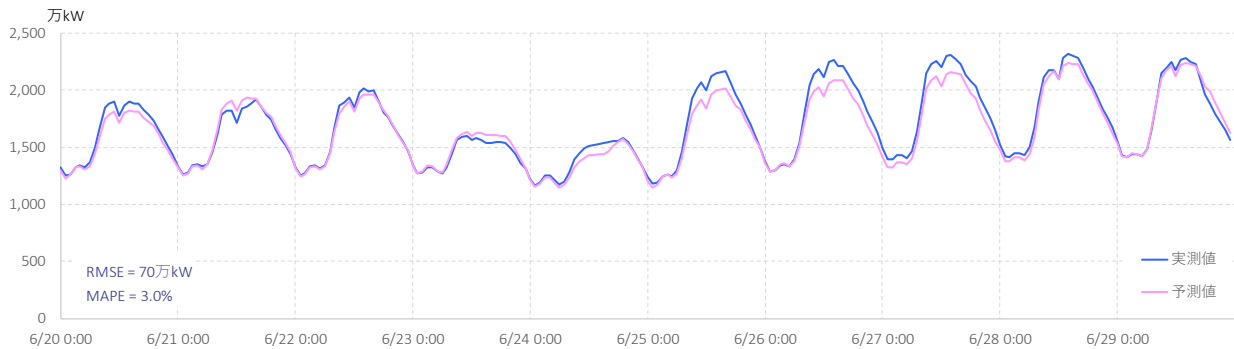


図 4-13 関西電力管内の予測結果例

(7) 中国

中国電力管内の6月20日～29日の平均電力需要は679万kW、最大は28(木)午後2時の876万kW、最小は24(日)午前1時の506万kWであった。24(日)～29(金)は各日の最大電力量に対してやや過小な予測が6日間連続で発生し、MAPEは10地域のうち8番目の精度となる2.8%となっている。この予測誤差は、最高気温の推移が25(月)に31.1°C(対前日+4.0°C)と2012年以降の当該日では最も高くなる一方、翌日の26(火)には26.1°C(対前日-5.0°C)と最高気温が急低下したことに加え、26(火)・28(木)は最高気温の予報との誤差がそれぞれ+2.9°C・-2.7°C、24(日)・29(金)は最低気温の予報との誤差がそれぞれ-2.2°C・+2.4°Cとなるなど、最高気温の乱降下や予報の誤差に応じて生じている可能性が考えられる。

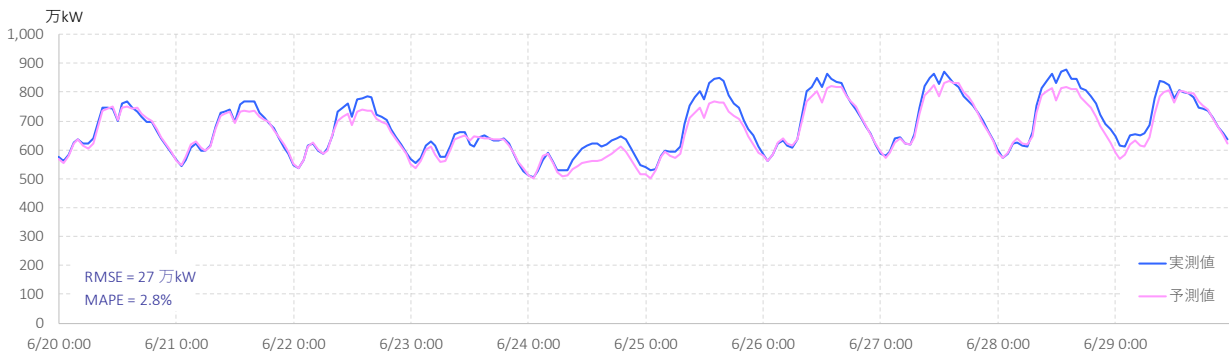


図 4-14 中国電力管内の予測結果例

(8) 四国

四国電力管内の6月20日～29日の平均電力需要は313万kW、最大は28(木)午後2時の426万kW、最小は24(日)午前7時の223万kWであった。20(水)・25(月)・26(火)・27(水)はやや過小な予測となっている。20(水)は最高気温23.7°Cに対して予想最高気温は25°Cと、天気予報の誤差に応じて予測誤差が生じたことが伺える。また、25(月)・26(火)・27(水)は2012～2017年の当該日における平均的な最高気温より3～5°C程度高く、例年よりも冷房需要が早めに現れたことにより予測誤差が生じている可能性が考えられる。

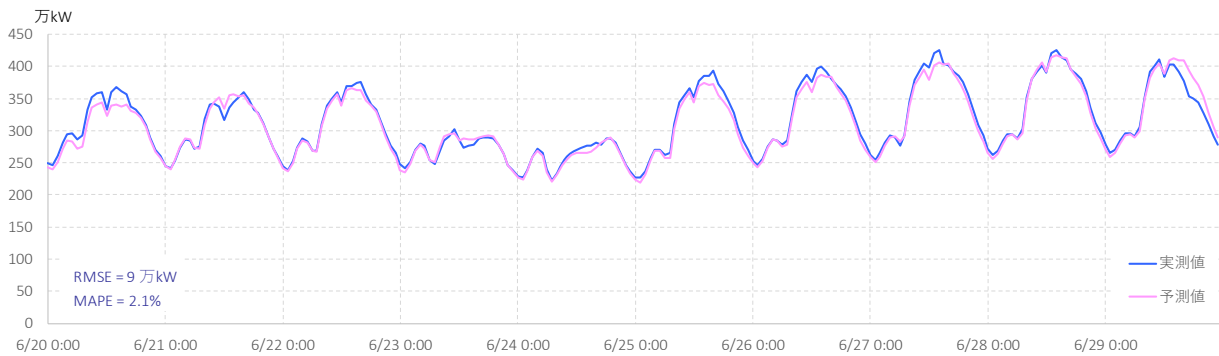


図 4-15 四国電力管内の予測結果例

(9) 九州

九州電力管内の6月20日～29日の平均電力需要は997万kW、最大は27(水)午後1時の1,273万kW、最小は24(日)午前1時の719万kWであった。22(金)～26(火)まで5日間連続で予測を大幅に外したことによりMAPEは10地域中最大の3.6%となった。そのうち、23(土)は予想最高気温31°Cに対して最高気温が24.5°C(誤差-6.5°C)であり、過大な予測をしたことが伺える。また、22(金)、23(土)、25(月)、26(火)には天気予報が外れて雨が降った。そうした不安定な気象の影響を受け予測精度が悪化した可能性も考えられる。

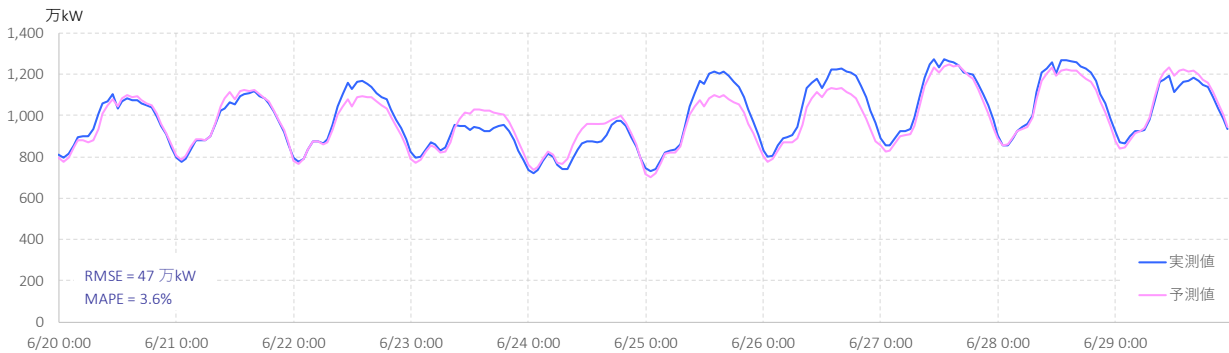


図 4-16 九州電力管内の予測結果例

(10) 沖縄

沖縄電力管内の6月20日～29日の平均電力需要は116万kW、最大は28(木)午後3時の141万kW、最小は23(土)午前3時の88万kWであった。同管内では、後述する年間平均の予測誤差がMAPE=3.0%と10地域中で最も大きいものに対して、当該期間の予測誤差はMAPE=1.7%と比較的良好な結果となっている。これは当該期間は電力需要の波形が各日とも似ており、安定的に推移していることに起因する可能性がある。一方、24(日)・25(月)はやや過小及び過大な予測となっている。但し、何れの日も天気予報による最高気温・最低気温の予測誤差は1℃未満であり、ここには大きな誤差は見受けられない。今回の評価では那覇のみの気象データを用いて予測を行っているが、沖縄のような多くの島嶼からなる地域の電力需要について、より多地点のデータを用いることによりどの程度まで予測性能が改善するかを探ることは興味深い課題である。

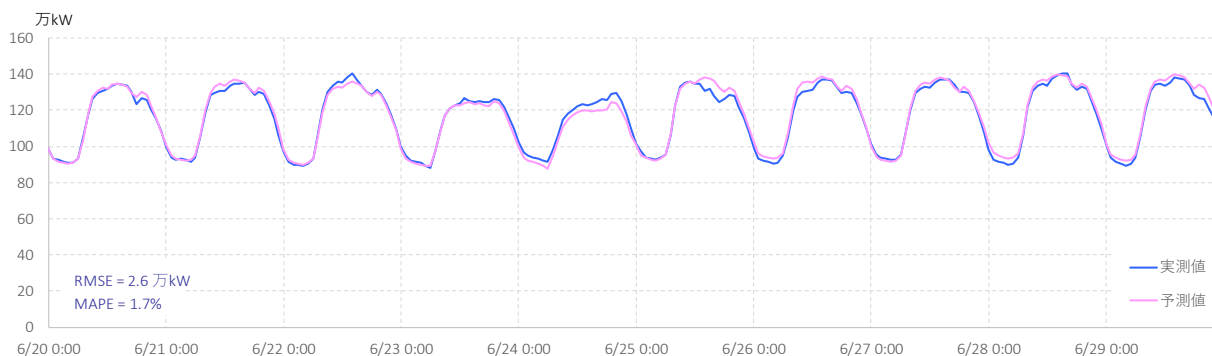


図 4-17 沖縄電力管内の予測結果例

4-2-2 予測月による誤差の変化

10地域について予測の誤差(MAPE)を月ごとに示すと表4-1及び図4-18の通りとなる。ここに示すように、多くの地域において夏(7月と8月)と冬(1月と2月)に誤差が大きい。地域別に見ると東北が年間平均で1.4%と著しく誤差が小さい一方で、中国や九州、沖縄において誤差が大きくなっている。但し沖縄の誤差が大きい理由の少なくとも一部は、沖縄のみ学習データが少ないことに起因すると考えられる。また図4-19に示す通り、予測誤差は日中に大きく、夜間に小さい。但し北海道や沖縄では、この形状は多地域とはやや異なるものとなっている。

北海道電力、北陸電力、関西電力及び九州電力については、各社のウェブサイト上にそれぞれの供給区域内における過去の電力需要の予測値が掲載されている。ここで対象とした2017年7月から2018年6月までの期間では、各社の予測におけるMAPEはおよそ3.1%、2.4%、2.5%及び2.4%であった。但しこれらの予測値は前日夜、もしくは当日朝など異なる時点の予測を含むために直接比較することはできない上に、電力会社は多くの場合、電力の供給安定性の観点から、需要を意図的に高めに想定する傾向にある¹²⁾ことにも注意が必要である。一方で本稿での予測は前日朝の時点想定しており電力会社の公表する需要予測時点よりも早いものであること、またごく限られた入力データのみに基づいていることなどを考えると、表4-1に見られる平均2.5%弱のMAPEは概ね良好な結果を示しているとも見られる。

表 4-1 地域別・月別の平均誤差率 (MAPE)

	2017年						2018年						年間平均
	7月	8月	9月	10月	11月	12月	1月	2月	3月	4月	5月	6月	
北海道	2.8%	2.1%	2.3%	2.2%	2.4%	2.5%	2.5%	2.3%	3.0%	2.6%	2.1%	1.8%	2.4%
東北	2.0%	1.8%	1.4%	1.0%	1.2%	1.5%	1.6%	1.9%	1.7%	1.1%	0.9%	1.2%	1.4%
東京	3.4%	3.9%	1.9%	1.9%	1.8%	2.0%	2.6%	2.8%	2.2%	1.6%	1.8%	2.1%	2.3%
中部	2.2%	3.8%	2.6%	2.2%	1.7%	2.2%	2.9%	1.9%	2.0%	2.0%	2.1%	1.6%	2.3%
北陸	3.0%	3.0%	2.1%	1.6%	2.5%	2.7%	3.0%	2.9%	2.6%	2.0%	2.2%	2.2%	2.5%
関西	2.6%	3.7%	2.4%	2.3%	1.9%	1.9%	2.5%	2.4%	2.1%	1.8%	1.7%	2.3%	2.3%
中国	3.3%	3.9%	3.0%	2.3%	2.3%	2.5%	3.0%	2.5%	2.6%	2.2%	2.8%	2.4%	2.7%
四国	3.0%	3.0%	2.6%	2.3%	2.0%	2.7%	3.1%	2.8%	2.5%	2.0%	1.9%	1.9%	2.5%
九州	2.7%	3.8%	3.3%	2.6%	2.0%	2.5%	3.1%	2.9%	2.6%	2.4%	2.0%	2.7%	2.7%
沖縄	2.3%	2.4%	3.8%	3.9%	2.7%	2.3%	2.8%	3.6%	2.4%	2.6%	3.6%	3.2%	3.0%

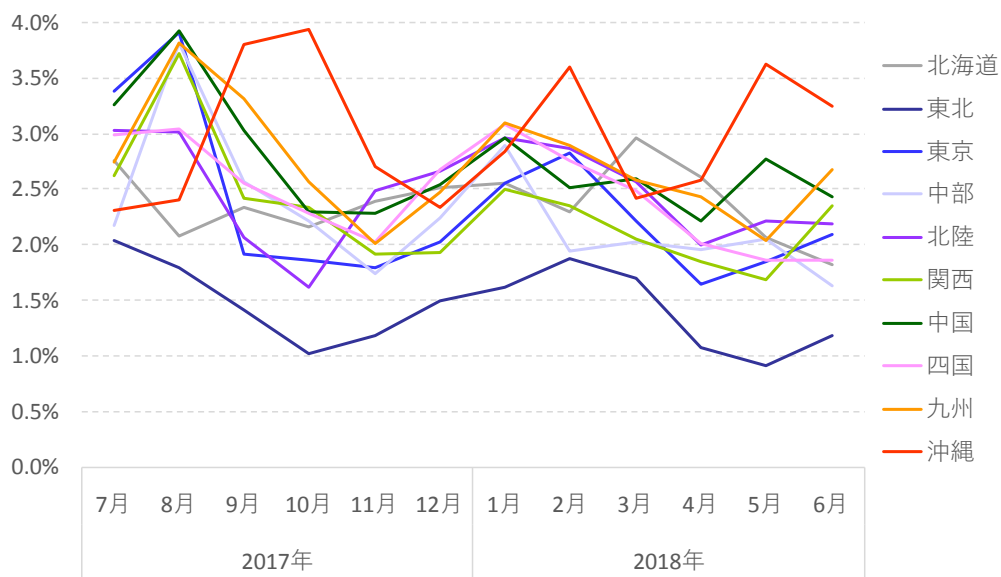


図 4-18 地域別・月別の平均誤差率 (MAPE)

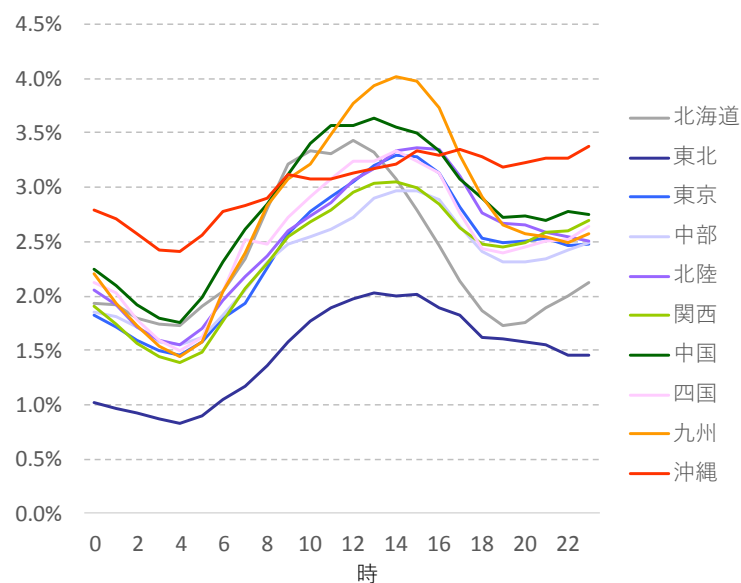


図 4-19 地域別・時間別の平均誤差率 (MAPE)

図4-20に地域別の予測誤差率（予測値と実績値との差を実績値で除した値）の分布を示す。このグラフには10地域それぞれについて、365日×24時間×15回試行（131,400データ）の誤差率を0.1%の位で四捨五入し、頻度として示している。誤差の分布はこのように沖縄で広く、東北で狭くなっているが、多くの地域で概ね誤差率±5%の範囲の中に位置しており、稀に10%に近い、もしくはそれを越える誤差を示す結果となっている。

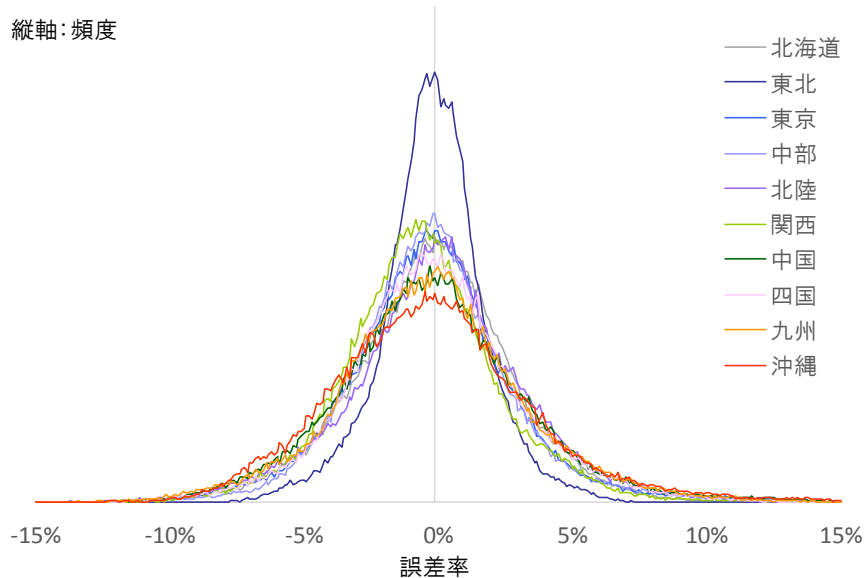


図 4-20 各地域の誤差率の分布

予測誤差が大きくなる夏季や冬季は、電力需要の変動が大きくなる時期であることは注意に値する。図4-21に示すように、横軸に各月内の電力需要の標準偏差を年平均需要で除した値、縦軸に各月のMAPEをとると、両者が緩やかな相関を示していることが伺える。このことから、例えば東北で予測誤差が小さい理由の一端は、附録2に示すように電力需要の変動が他の地域に比べて小さいことにあると見ることもできるであろう。また北海道において他の地域と異なり1月や2月よりもむしろ3月に誤差が大きくなっていることは、北海道の3月の需要の標準偏差が1月や2月よりも大きくなっていることと関係しているかも知れない。但し需要の標準偏差のみで予測誤差の大小が決るわけでは勿論なく、例えば中国で5月に誤差が大きくなっている理由は不明である。

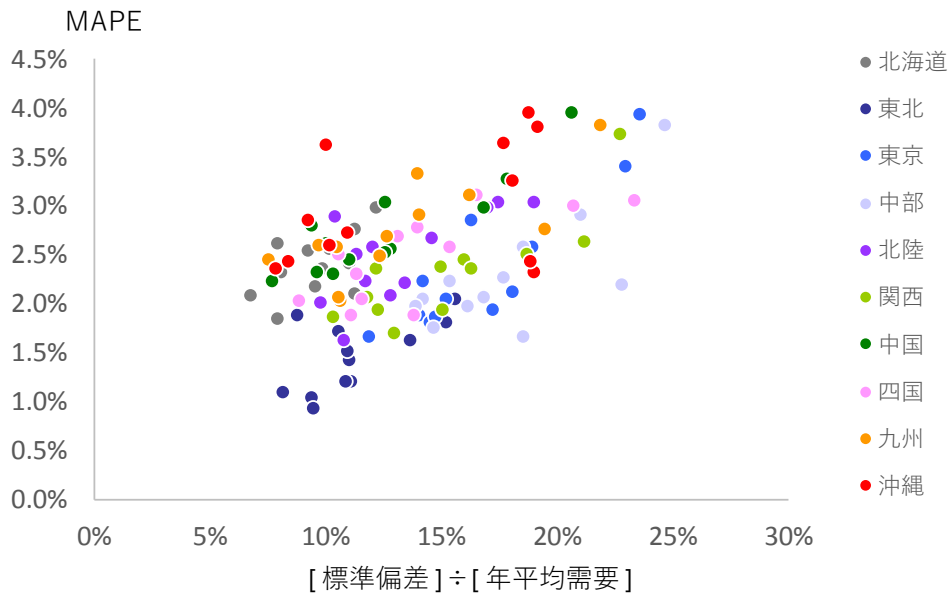


図 4-21 電力需要の標準偏差と MAPE との関係

5. まとめ

本稿では国内の電力 10 供給区域を対象とし、1 地点の気象データのみを利用する簡易な人工ニューラルネットワークを用いて短期電力予測を試みた。ここでは需要予測のタスクを 3 つのステップに分割し、各ステップにおいて選択的アンサンブル平均をとることで、より正確な予測を行うことが可能となっている。但し予測性能は季節や地域によって異なり、電力需要の変動が大きな夏季及び冬季には誤差が大きく、春季及び秋季には比較的小さな誤差となる。また地域別にみると東北で誤差が小さく、中国や九州では比較的大きいなどの差もあり、更に北海道や沖縄では他の地域と異なる月に誤差が大きくなるなど、地域による特性も現れている。

今回用いたモデルに対し、湿度や 1 時間ごとの気温予報値等、より多種類の入力データを用いることや、より多地点の気象データを用いること、選択的アンサンブルの試行回数を増やすことなどで更に予測性能を向上させることは可能であると考えられる。但し本稿に示されるように、ある特定のデータセットの上で良好な予測性能を示すモデルが、常に良い結果を示すとは限らない。これは、予測モデルの優劣を比較するためには慎重な配慮が必要であり、可能な限り大きなデータの上で評価を行うことが望ましい、ということを示している。またモデルによって異なる季節に得手・不得手があると考えられることから、例えば本手法を含む ANN による予測モデルを他の予測手法と組み合わせることで精度を向上させ得る可能性もある。一方で例えば選択的アンサンブル平均の手法は予測性能を向上させるために有効であり、他の予測問題に対しても応用可能であることが期待される。

但し本手法の延長でどの程度まで予測性能を向上させたとしても、夏季や冬季には依然として比較的大きな誤差を示すものと思われる。大きな予測誤差が生じる日は比較的少数であり、その誤差のかなりの部分は天気予報そのものの誤差から生じているものの³、そのみが誤差の要因でないことも本稿で示す結果からは推測される。誤差が生じる要因を更に追求し、改善を図ることは今後の重要な研究課題である。

³ なお天気予報の誤差は完全にランダムに生じるものではなく、ある程度の傾向性をもつことが過去のデータからは推測される。これを明示的に考慮することにより、少なくともある程度は予測性能を向上させ得る可能性が高い。

参考文献

- 1) M. Tucci et al., 2016. A multi-objective method for short-term load forecasting in European countries. *IEEE Trans. Power Syst.*, 31(5), pp. 3537-3547.
- 2) TESLA, Inc. "Our Method." <http://www.teslaforecast.com/about-tesla/our-method/>
- 3) 経済産業省, 2017. 「第 25 回制度設計専門家会合事務局提出資料」
http://www.emsc.meti.go.jp/activity/emsc_system/pdf/025_05_00.pdf
- 4) 松原雄樹, 2018. 「供給エリア内の電力需要の予測システムと太陽光発電の出力予測」, 『エネルギー・資源』, 39(1), pp.50-54.
- 5) K. Metaxiotis et al., 2003. Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher. *Energy Convers. Manag.*, 44(9), pp. 1525-1534.
- 6) F. M. Bianchi et al., 2017. *Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis*. Springer. ISBN: 978-3319703374.
- 7) T. Hong et al., 2014. Global energy forecasting competition 2012. *Int. J. Forecasting*, 30(2), pp. 357-363.
- 8) T. Hong et al., 2016. Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond. *Int. J. Forecasting*, 32(3), pp. 896-913.
- 9) Réseau de transport d'électricité, 2018. Prévision déterministe de la consommation électrique hivernale (volet 1). <https://www.datascience.net/fr/challenge/33/details>
- 10) 東京電力ホールディングス株式会社, 2017. 「電力需要予測値の正確さを競う「第 1 回電力需要予測コンテスト」の開催について」 http://www.tepco.co.jp/press/news/2017/1440911_8963.html
- 11) R. Salkuti. (2018). Short-term electrical load forecasting using radial basis function neural networks considering weather factors. *Electr. Eng.* pp.1-11. <https://doi.org/10.1007/s00202-018-0678-8>
- 12) 北海道電力株式会社, 2017. 「今後の需要想定方法について」
http://www.emsc.meti.go.jp/activity/emsc_system/pdf/025_05_01.pdf
- 13) D. P. Kingma and J. Ba, 2014. Adam: A method for stochastic optimization. <https://arxiv.org/abs/1412.6980>
- 14) Y. Matsuo and T. Oyama, 2018. Forecasting daily electricity demand by applying artificial neural network with Fourier transform and principal component analysis techniques. The 13th International Symposium on Operations Research and its Applications, Session 5-3.
- 15) 松尾雄司, 大山達雄, 2018. Short-term electric demand forecasting using artificial neural networks with daily load curve analysis techniques. 日本オペレーションズ・リサーチ学会 2018 年秋季研究発表会, 2-E-9.
- 16) Z.-H. Zhou et al., 2002. Ensembling neural networks: Many could be better than all. *Artif. Intell.*, 137, pp. 239-263.

附録1 主成分分析と選択的アンサンブル平均法を用いた電力需要予測モデル

附1-1 人工ニューラルネットワーク (Artificial Neural Network: ANN)

一般的に ANN は図2-1 に示される通り、入力層・出力層と1つ、もしくは複数の中間層によって構成される。いま入力データが N_0 次元のベクトル $\mathbf{y}_0 = \mathbf{X}$ で表され、 n 番目の中間層が N_0 次元のベクトル \mathbf{y}_n で表されるとすると、各層の間に次の形の関係式を想定することにより、入力データ \mathbf{X} から出力データ $\mathbf{y}_{out} = \mathbf{y}_{M+1}$ が逐次的に計算されることになる。

$$\mathbf{y}_{n+1} = \phi(\mathbf{w}_n \mathbf{y}_n + \mathbf{b}_n) \quad n \in \{0, 1, \dots, M\} \quad (A1)$$

ここで ϕ は活性化関数と呼ばれる非線形の関数であり、行列 \mathbf{w}_n とベクトル \mathbf{b}_n はウェイト及びバイアスと呼ばれるパラメータである。この活性化関数の非線形性によって複雑な事象をモデル化し、精度の良い予測を行うことが可能となる。

ニューラルネットワークの学習に際しては、学習データとして大量の入力データ \mathbf{X} と出力データ(教師データ) \mathbf{Y} の組を用意する。そして、入力データをネットワークに与え、得られた出力データと教師データとの誤差(多くの場合、 \mathbf{y}_{out} と \mathbf{Y} の間の二乗ユークリッド距離が用いられる)を最小にするように、 \mathbf{w}_n と \mathbf{b}_n が最適化される。ここでは勾配降下法によってこの最適問題を解くことが一般的であり、より効率的に解を求めるために、勾配降下法のさまざまなバリエーションが提案されている。いずれにせよ、勾配降下法は最初に乱数によって \mathbf{w}_n と \mathbf{b}_n の初期値を定め、そこから最適な \mathbf{w}_n と \mathbf{b}_n を目指して勾配を降下してゆくものであり、初期値の与え方によって結果がある程度異なるものとなることに注意する必要がある。中間層の数 M と各層のニューロン数 N_i はモデルを規定するパラメータであるが、これらを設定する明確な指標は存在せず、課題に応じて試行錯誤的に設定する必要がある。本稿で用いたモデルでは ϕ としてはソフトプラス関数 $\phi(x) = \log(1 + e^x)$ を用い、また層数 M は3、各層のニューロン数 N_i は30としている。また勾配降下法としては Adam¹³⁾ を用いた。

附1-2 主成分分析による日負荷曲線の解析

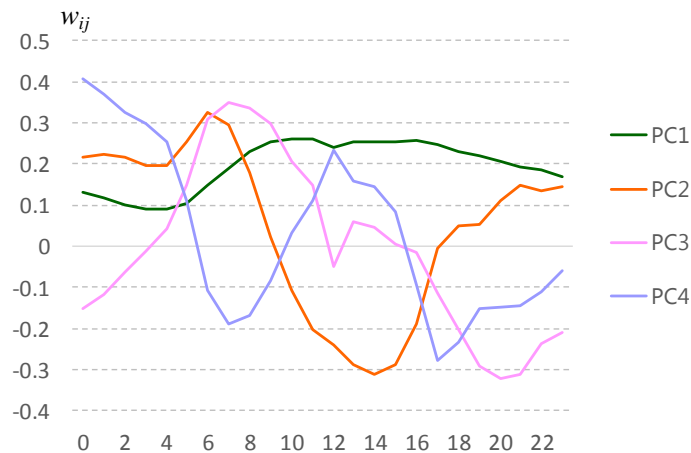
主成分分析は多数のデータから主要な特徴を抽出するために広く用いられる手法である。いま m 個の変数 x_1, x_2, \dots, x_m をもつデータセットが与えられたとする(即ち、次元数 m の n 組のデータ $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n, \mathbf{X}_k = [x_1^{(k)}, x_2^{(k)}, \dots, x_m^{(k)}], k \in [1, 2, \dots, n]$ が与えられたとする)。これらの変数の線形結合

$$z_1 = w_{11}x_1 + w_{21}x_2 + \dots + w_{m1}x_m \quad (A2)$$

について、 $w_{11}^2 + w_{21}^2 + \dots + w_{m1}^2 = 1$ の条件のもとで z_1 の分散が最大となるように係数 $w_{i1} (i=1, 2, \dots, m)$ が与えられたとき、この z_1 を第1主成分と呼ぶ。即ち、変数 x_1, x_2, \dots, x_m の線形結合として与えられる新たな変数の中で、このデータセットの特徴を最もよく表現できるように選ばれたものが第1主成分 z_1 である。同様に、 z_1 と直交する $m-1$ 次元の空間の中で分散が最も大きくなるように選ばれた $z_2 = w_{12}x_1 + w_{22}x_2 + \dots + w_{m2}x_m$ (但し $w_{12}^2 + w_{22}^2 + \dots + w_{m2}^2 = 1$) が第2主成分と呼ばれる。この操作を続けることにより第 i 主成分 ($i < m$) までの変数を作成し、それらによって与えられたデータセットを十分に良く近似できたならば、 m 次元のデータを i 次元に縮約することに成功したことになる。数学的には各主成分は共分散行列の固有ベクトルとして求めることができる。

ある日の電力需要(1時間平均値)は24次元空間内のベクトルと見做すことができる ($m=24$)。例として東京電力管内の負荷曲線について主成分分析を行い、第1主成分(PC1)から第4主成分(PC4)までを図示すると図附-1のようになり、ある特定の日の負荷曲線はこれらの主成分の線形結合として近似することが可能となる。

即ちこの場合、24次元の電力需要データは4つの各主成分に対応する係数（主成分得点と呼ばれる）によって近似的に表される。本稿で用いたモデルでは、予め過去の全ての日について4つの主成分得点を算出しておき、それを学習に用いることで、翌日の4つの主成分得点を予測する。この次元の圧縮によって有意に予測性能を高めることが可能となる^{14,15)}。



図附-1 主成分分析の例（東京電力管内）

なお機械学習の分野では、自己符号化器（Autoencoder）を用いることによって主成分分析を更に一般化した次元圧縮が可能となることが知られている。このため主成分分析の代りに自己符号化器を用いることによって、本稿で用いたモデルと同等以上の性能を有するモデルを作成することは可能であると思われる。但し電力需要の分析に限って言えば第4主成分までの累計寄与率が99%以上となるため、簡単な主成分分析によってかなり正確な近似が可能であること、また主成分得点値を予測することにより、需要の形状の変化をある程度明示的に把握できることなどから、今回のモデルでは主成分分析を用いている。

附 1-3 選択的アンサンブル平均法

機械学習による予測を行う際、ただ一つのモデルの結果を採用するのではなく、複数のモデルの結果の平均値を取ることで予測性能が向上することが広く認識されている。またこの特性を利用し、複数のモデルに効率的に学習を行わせ、その結果を利用して精度の良い予測を行うことを目指す「アンサンブル学習法」（より具体的には手法に応じてバギング、ブースティングなどと呼ばれる）が多くの研究で用いられている。

これに対し、Zhou et al. (2002)¹⁶⁾は「選択的アンサンブル平均」による予測性能、もしくは分類性能の向上の可能性を示した。これは複数のモデルによる結果をそのまま全て平均して予測値を作成するのではなく、まず予測以前の検証データにおけるモデル出力の誤差を評価し、そこで誤差の大きいモデルは除外して、残りのモデルのみでアンサンブル平均を取るものである（Zhou らはここで除外するモデルを選択するために遺伝的アルゴリズムを用いている）。

本稿で用いた手法は検証データにおけるモデルの誤差が予測データ（テストデータ）における予測誤差と有意に相関するように検証データを設定し、その上でモデルの選択を行うものであり、「経験的」選択アンサンブル平均法 (Empirical Selective Ensembling) と呼ぶべきものである。例えば40人の生徒がいる学級を想像されたい。Zhou らの手法が出来る悪い数人の生徒を除外し、残りの多くの生徒の回答の平均値をもって最終的な回答とする方法であるのに対し、ここで用いた手法は出来る良い数人のみの回答の平均をとって最終的な回答とするものであり、そのために生徒の出来る良い/悪いを効率的に判断できるよう、予め判断基準を経験的に準備しておくものであると言える。

これはここで対象とする課題 (STLF) における以下の観察に基づいている。即ち多くの機械学習モデルにおいて用いられているように、学習データの中からランダムに検証データを選択した場合、検証データとテストデータとの誤差の間に有意な相関はほとんど見られない。これに対し、テストデータの直前の数日を検証データとした場合、両者の誤差には多くの場合有意な相関を見ることができる。これは STLF における学習期間と予測期間との長さのギャップによるものと考えられる。つまりいま予測したいデータはわずかに 1 日先の電力需要であり、その需要の変化のあり方は数年前のものとは大きく異なっていると思われる。一般的に深層学習を行う際にはより多くの学習データを用いた場合の方がより高い汎化性能 (予測性能) を期待でき、実際に本稿で示した通り、短い学習期間のみから学習を行うよりもより遠い過去からの多くのデータを用いて学習した方が予測性能が良くなることが確認できるものの、明日の予測を行う場合に 3 日前のデータと 3 年前のデータが等しく参考になるわけではない、ということも同時に容易に想像されるであろう。

具体的な手順は以下の通りである。例えば 2018 年 1 月の電力需要をテストデータとして予測性能を評価する場合、その直前の n_v 日 (例えば $n_v=10$ であれば 2017 年 12 月 22 日~31 日) を検証データとし、それ以前 (2017 年 12 月 21 日まで) を学習データとする。この学習データを用いて n_T 組の初期条件からモデルを学習させ、得られた n_T 個のモデルに対し、それぞれ検証データ上での誤差を評価する。その誤差が小さい方から順に n_S 個 ($1 \leq n_S \leq n_T$) のモデルを選択し、それらの出力の算術平均値による誤差が検証データ上で最小となるように、 n_S の値を設定する。最後に、 n_S 個のモデルの予測値の算術平均値をもって、最終的な予測値とする。

経験的選択アンサンブル平均の名の通り、 n_v の値は経験的に設定される。一般的には n_v が小さいと評価が安定しなくなる一方で、 n_v が長すぎるとテストデータと相関の少ない日の値まで取り込んでしまい、選択の効果が薄くなるものと考えられる。ここでは $n_v=30$ とした。また n_T は当然ながら大きい方が (40 人のクラスから優秀な生徒を選ぶよりも 200 人の学年全体の中から優秀な生徒を選んだ方が) 予測性能が向上すると考えられ、実際にそのような傾向を確認することができる。但しこれは計算時間とのトレードオフになる。本稿では一律に $n_T=20$ として計算を行っている。

附 1-4 短期電力需要予測モデルの概要

本稿では、主成分分析と選択的アンサンブル平均法を用いた短期電力需要予測モデルを用いて評価を行っている。ここでは d 日の朝 8 時に $d+1$ 日の 24 時間値を予測することを想定し、以下のデータをモデルへの入力データとする。

電力需要データ： $d-1$ 日 (24 時間値)、 d 日 (午前 7 時まで)

暦データ (年、月、日、曜日、祝日)： $d-1$ 日、 d 日、 $d+1$ 日、 $d+2$ 日

気象データ (最高気温、最低気温、天候)： $d-1$ 日

気象予報データ (最高気温、最適気温、天候)： $d-1$ 日、 d 日、 $d+1$ 日

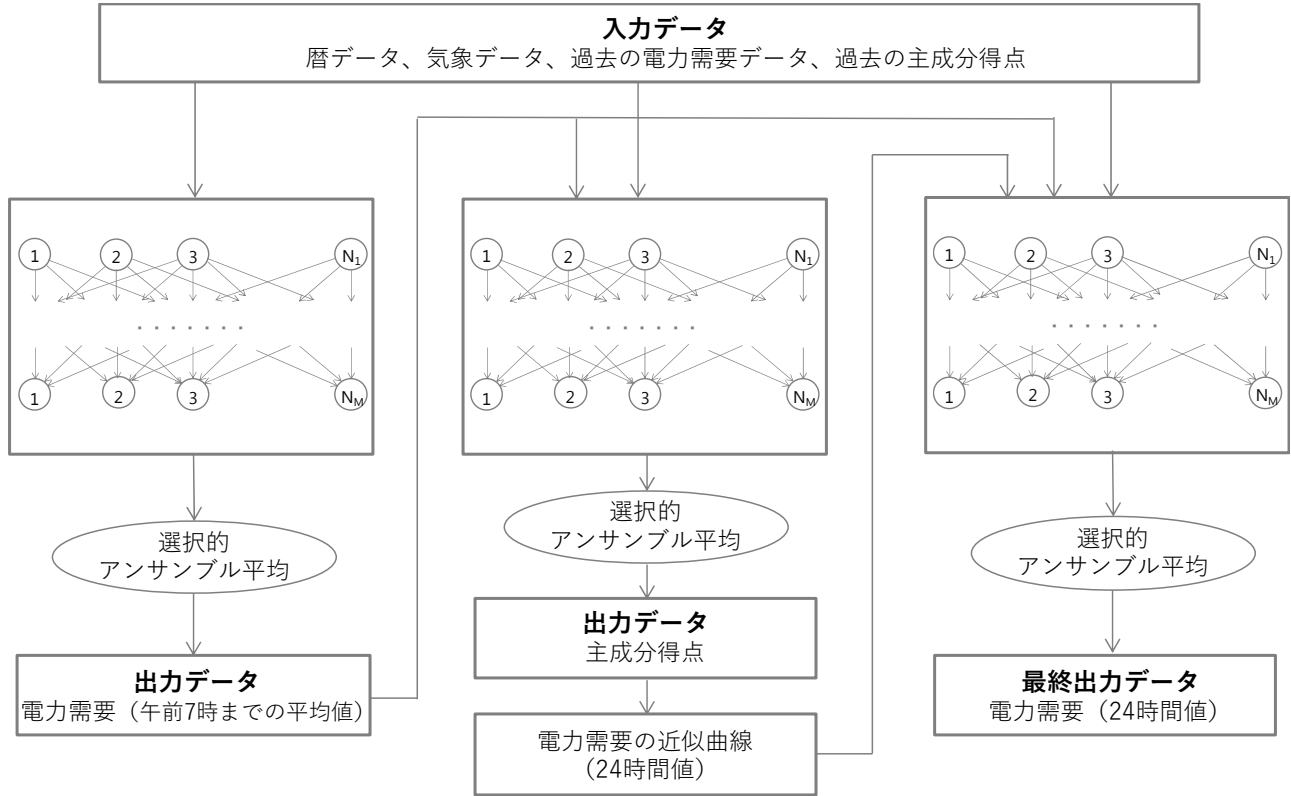
暦データについては年、月、日はその自然数値そのものを用い、曜日については日曜を 0、月曜を 1、・・・、土曜を 7 とする。また祝日については、その日が国民の祝日もしくは 12 月 29 日~1 月 3 日、8 月 13 日~16 日の期間内であれば 1、それ以外であれば 0 とする。気象及び気象予報データについては、日最高気温及び最低気温は°C 単位の値をそのまま用い、また天候については晴 (W_1)、雨 (W_2)、曇 (W_3)、雪 (W_4) の 4 つの変数を用意し、気象庁の公表する天気概況にそれぞれの語が含まれていれば 1、いなければ 0 とする。例えばその日の天気概況が「晴時々曇、一時雪、ふぶきを伴う」であった場合には、 $W_1=W_3=W_4=1$ 、 $W_2=0$ となる。気象予報については、前日朝時点のものを用いている。更に全ての値を次式に従って規格化し、モデルの入出力データとし

た。

$$x_t = \frac{X_t - \bar{X}}{X_{max} - X_{min}} \tag{A3}$$

ここで X_t はもとの値、 \bar{X} 及び X_{max} 、 X_{min} はそれぞれ X_t の平均値、最大値及び最小値、 x_t はモデルの入出力データの値である。

モデルの概要は図附-2 に示す通りである。



図附-2 短期電力需要予測モデルの概念図

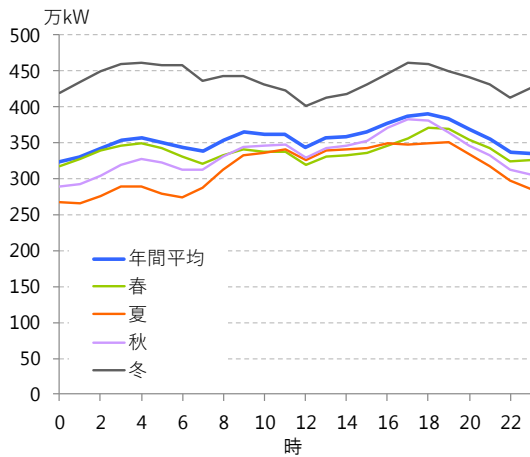
ここに示す通り、本モデルではタスクを3回に分割して予測を実施している。まず第一ステップでは、 $d+1$ 日の午前0時~7時の電力需要平均値 $h_{7,d+1}$ を予測する。これは他の値が $d-1$ 日までしか実績値がないのに対し、この午前7時までの平均値は d 日までの実績値が存在し、より正確な予測が可能となるためである。第二ステップではこの $h_{7,d+1}$ を入力データに加え、また $d-1$ 日までの主成分得点実績値 (第4主成分まで) も入力データとした上で、 $d+1$ 日の主成分得点を予測する。この予測された主成分得点から $d+1$ 日の電力需要の近似曲線を得た上で、第三ステップでそれをも入力データに加えて $d+1$ 日の電力需要の24時間値を予測する。なお各ステップにおいて $n_T=20$ として選択的アンサンブル平均を求め、それを予測値とした。なお平均的に電力需要の予測値が実測値よりもやや小さめに出るバイアスが見られたため、予測日から遡って30日程度の平均予測誤差 (24時間ごと) を評価し、その値を最終出力データから控除した。この補正により、多くの地域で MAPE が平均して0.1%程度改善している。

電力負荷曲線を分析する手法としては、主成分分析の他にも離散 Fourier 変換や k 平均法、ファジー c 平均法など様々な方法が存在する。より単純には、主成分得点の代りに電力需要の日平均値、午前平均値、午後平均値などを予測し、それらを入力データとして24時間値を予測する方法も考えられる。筆者らが試みた範囲内では離

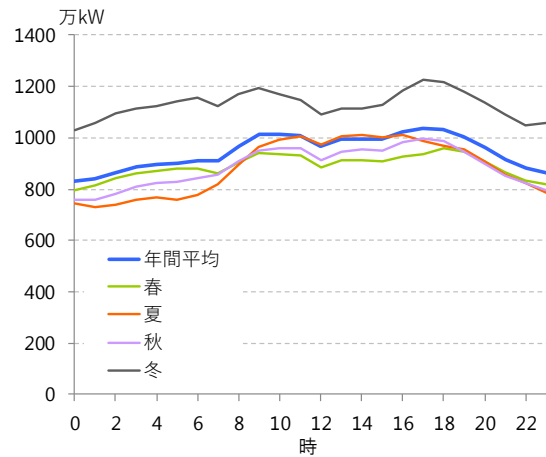
散 Fourier 変換や午前平均値・午後平均値等を使う方法でも主成分分析と概ね同等、もしくはやや劣る予測結果が得られる一方で、ファジーc 平均ではより精度の悪い結果しか得られなかった。但し二つの異なるモデルを比べる場合、それぞれのモデルによって予測精度が良くなる季節（月）や悪くなる季節（月）が若干異なるため、モデル間の優劣を決めることは一般的には容易でない。季節や電力負荷曲線の形状と各種モデル化手法の適性とを評価することは今後の興味深い課題である。

附録2 各地域内の電力負荷曲線

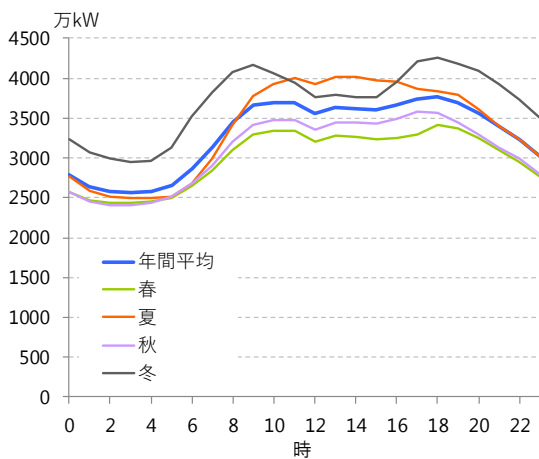
図附-3 に各供給区域における春（3～5月）、夏（6～8月）、秋（9～11月）、冬（12～1月）及び年間平均の日負荷曲線を示す。これは2017年4月1日から2018年3月31日までの電力需要値について、各季節内で平均をとったものである。多くの地域で夏・冬に需要が大きく、春・秋に小さいが、北海道では夏に小さく、沖縄では冬に小さいなど地域による特性が大きく現れている。また春と秋で比べると春の方が小さい。これは部分的には、同じ気温であっても近い過去の記憶によって冷房のつけ方が異なるという人間の特性に基づくものと思われる。



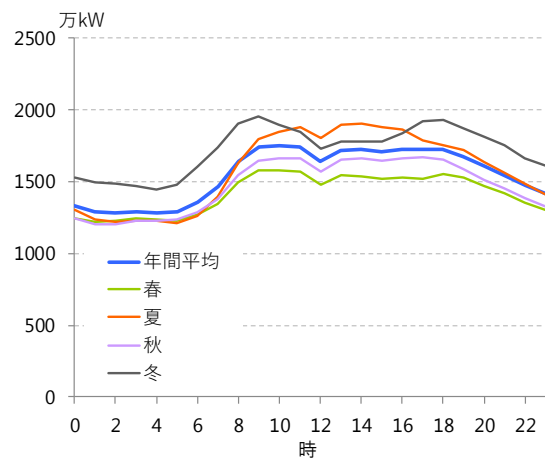
北海道



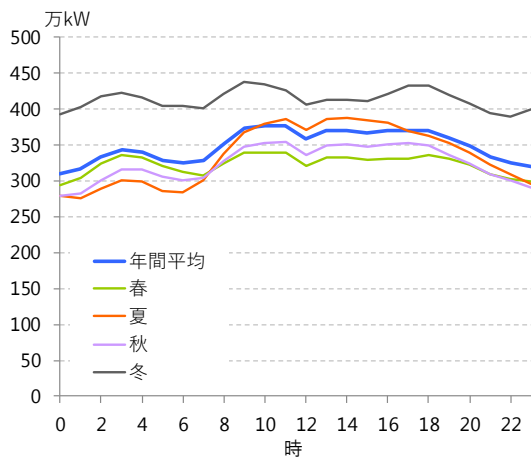
東北



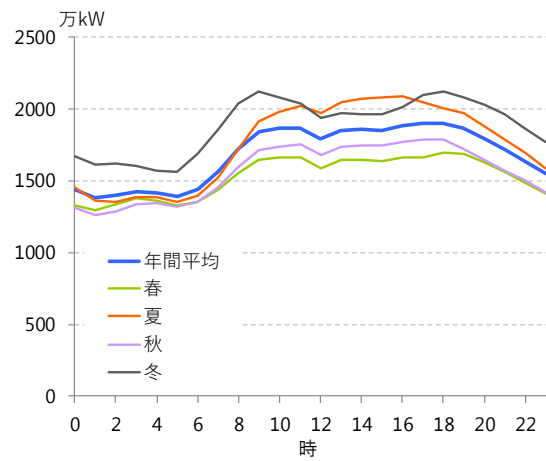
東京



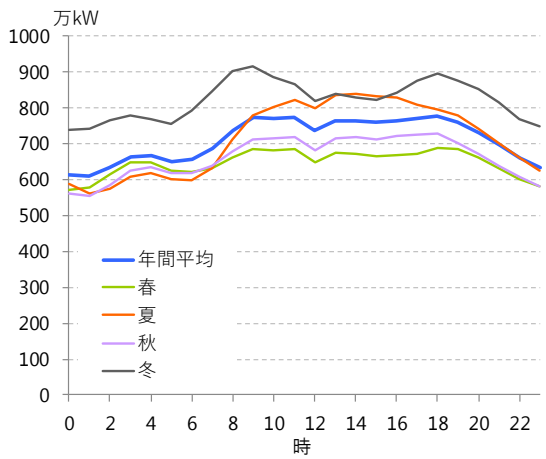
中部



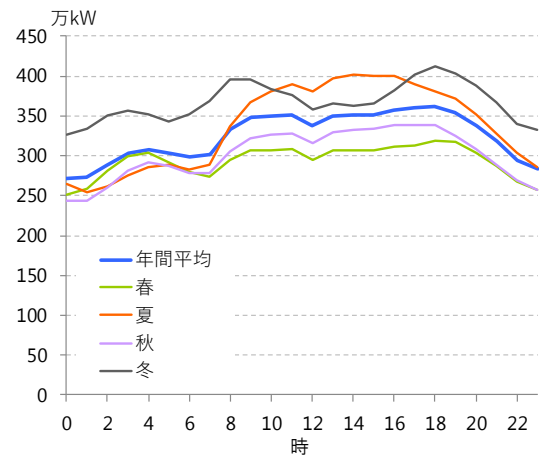
北陸



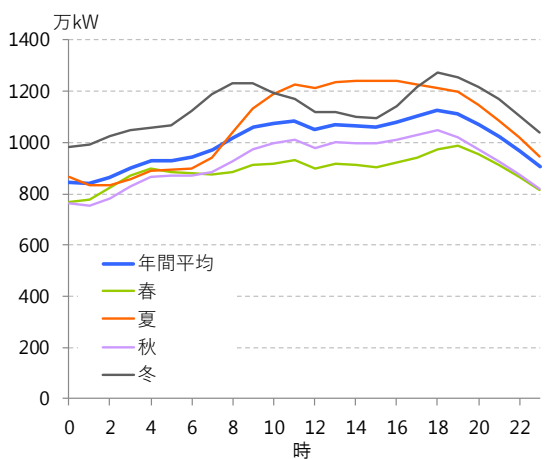
関西



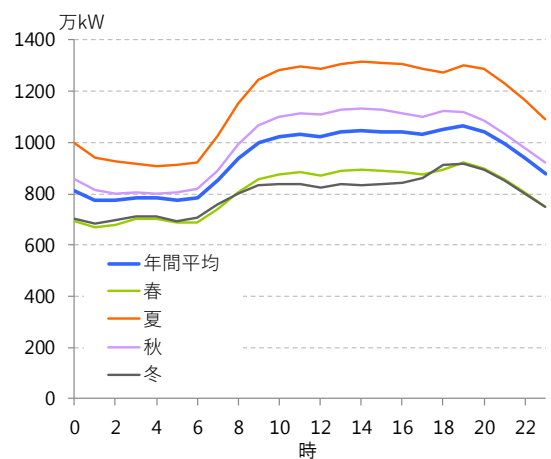
中国



四国



九州



沖縄

(出所) 各社HP、公表資料より作成

図附-3 各地域の季節別平均日負荷曲線 (2017年4月~2018年3月)