

電力データを用いた時系列予測手法の精度比較

關谷 純一[†]村上 晴美[‡]大阪公立大学大学院情報学研究所[†]トランスコスモス株式会社[‡]

1. はじめに

我々は、時系列予測手法の選定のため、代表的な手法として統計、機械学習の手法をバランスよく採用し、シンプルで標準的な実験の枠組みを作ることを提案している。先行研究（前回の実験）では、M4 Competition [1] の時系列データを用いて、M4 Competition でベンチマークとして採用された手法を基本に、機械学習の競技大会でよく使用される勾配ブースティング等を追加した 31 の時系列予測手法の精度比較実験を行った[2]。M4 ベンチマークに勾配ブースティング等の予測手法を加えることの有効性を確認した。

本研究では、新たに 2 つの統計手法、1 つの機械学習手法、2 つの複数系列の機械学習手法を追加し、電力データを用いて実験を行った。本稿では、電力会社 10 社の日時の時系列データ ($\Delta T=1$ 時間) を使用し、36 の時系列予測手法の比較実験を行った結果を報告する。

2. 方法

先行研究の 31 の時系列予測手法に加えて、2 つの統計手法、1 つの機械学習手法、2 つの複数系列の機械学習手法を追加した 36 の手法で実験を行う。

まず、時系列の予測手法でよく利用される、統計手法の Prophet と SARIMA と機械学習手法の Lasso を追加する。追加後の予測手法の内訳は統計が 14 手法、機械学習が 14 手法と同等の数となり、機械学習手法が多かった先行研究と比べバランスが良くなっている。

さらに、M4 Competition では参加者が採用した複数系列の機械学習手法の予測精度が高かったため本研究でも採用を試みる。本研究では、Random Forest と DeepAR について、単系列の予測手法だけではなく、複数系列の予測手法の実験を行う。Random Forest を選択した理由は前回の実験において機械学習手法の中で最も精度が高かったからである。一方、DeepAR は最新の手法であり、他の機械学習手法と比較して実装が容易なため選択した。なお、統計の各手法は複数系列に対応していない。

M4 Competition では多様な業種の時系列データが含まれたが、本研究では、電力会社 10 社（北海道電力、東北電力、東京電力、中部電力、北陸電力、関西電力、中国電力、四国電力、九州

電力、沖縄電力）の電力使用実績の時系列データ ($\Delta T=1$ 時間) を使用する。

- 2021/12/01 00:00:00～2022/11/30 23:00:00 までの時系列データを学習データとして使用（1社あたり：N=9,336 $\Delta T=1$ 時間）
- 2022/12/01 00:00:00～2022/12/31 23:00:00 までの時系列データをテストデータとして使用（1社あたり：N=168 $\Delta T=1$ 時間）

単系列の予測手法では、電力会社ごとの 9,336 件の学習データで 10 モデルを構築、複数系列の予測手法では、電力会社 10 社合計 93,360 件の学習データで 1 モデルを構築し、テストデータを使用して予測した。

予測手法の精度比較では、先行研究と同様に sMAPE と MASE の 2 つの評価指標を用い OWA でランキングした。

$$OWA = \frac{\frac{sMAPE_a}{sMAPE_b} + \frac{MASE_a}{MASE_b}}{2}$$

ここで a は対象の予測手法、b はベースラインの予測手法となる。ベースラインの予測手法は Naive S を用いている。sMAPE と MASE の値が共に小さければ予測手法の精度は高く、値が大きければ精度は低い。そのため OWA の値が小さいほど精度が高いことになる。

3. 結果

sMAPE を横軸、MASE を縦軸にして各手法をプロットした(図 1)。番号は予測手法の Rank を示す。sMAPE と MASE の値がいずれも小さい予測手法は複数系列の DeepAR、コンビネーションの Comb2, Comb3, Comb5, Comb6 である。一方、sMAPE と MASE の値がいずれも大きい予測手法は RNN, MLP, Damped であり、他とかけ離れていることがわかる。コンビネーションの予測精度が高く、M4 ベンチマークの RNN, MLP の予測精度が低いことや、複数系列の DeepAR がコンビネーションの予測精度と遜色がなく、他の単系列の予測手法よりも予測精度が高いことがわかる。

OWA でのランキングを表 1 に示す。比較のために先行研究[2]での結果も示す。1 位の予測手法は 2 つの精度の高い機械学習手法の組み合わせである Comb5 であった。次に精度の高い予測手法は 3 つの精度の高い予測手法の組合せである Comb2 であった。ランキング上位はコンビネーションが占める結果となった。単一の予測手法よりもコンビネーションの予測手法の精度が高いのは M4 Competition や先行研究と同じである。

Comparison of accuracy of time series forecasting methods using electricity data

[†]Juniichi Sekitani, Harumi Murakami, Graduate School of Informatics, Osaka Metropolitan University

[‡]Juniichii Sekitani, transcocosmos inc.

コンビネーション以外で最も精度の高い予測手法は、新たに追加した複数系列の DeepAR であった。次に精度の高い予測手法は GBDT、複数系列の Random Forest であり、機械学習手法が上位に出現した。また複数系列の DeepAR と Random Forest と単系列の Random Forest は精度が高かったが、単系列の DeepAR は精度が低かった。複数系列の予測精度はどちらも高く、複数系列の予測手法を採用することの可能性を示した。

先行研究で機械学習手法の中で最も精度が高い手法は Random Forest であったが、今回も予測精度が高い。Random Forest はデータの種類によらず高い予測精度が出せると考える。本研究では単系列の機械学習手法では GBDT が最も精度が高い。

統計手法を確認すると最も予測精度の高い手法は Auto ARIMA であり、他の手法はすべて M4 ベースラインである Naïve 1 よりも精度が低かった。今回追加した Prophet は他のほとんどの統計手法よりも予測精度が低い結果となった。

M4 Competition および先行研究とは異なり、全体的に統計手法は機械学習手法よりも精度が低い結果となった。これはデータの種類によるものなのか、今後の追加実験で確認が必要である。

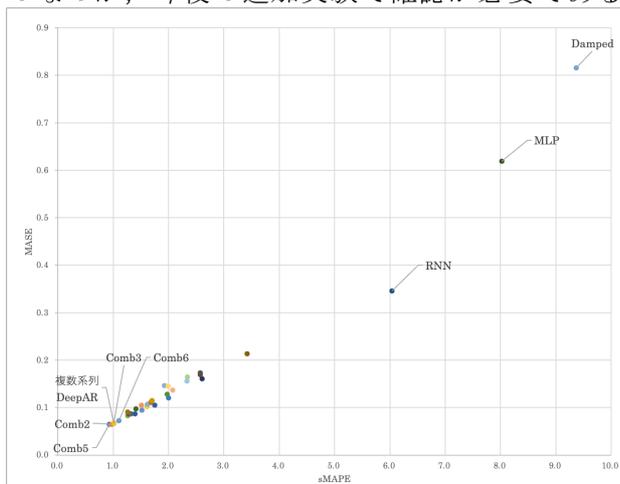


図 1：採用した 36 手法の sMAPE と MASE

4. おわりに

電力会社 10 社の日時時系列データ ($\Delta T=1$ 時間) を使用し、先行研究に Prophet や複数系列の DeepAR や Random Forest 等の予測手法を追加した 36 の予測手法で精度を比較した。

M4 Competition や先行研究と同様に上位はコンビネーションが占めた。コンビネーション以外では複数系列の DeepAR の予測精度が最も高かった。先行研究と異なり、全体的に機械学習手法が統計手法よりも精度が高かった。Random Forest は先行研究と同様に精度が高く、単系列より複数系列の精度が高かった。フレームワークに複数系列の予測手法を加えることが有効である可能性を示した。

表 1:36 手法のランキング

手法	種類	OWA	Rank	先行研究 [2] Rank
Comb5	コンビ	0.371	1	13
Comb2	コンビ	0.379	2	3
複数系列 DeepAR	機械学習	0.393	3	-
Comb3	コンビ	0.393	4	1
Comb6	コンビ	0.427	5	4
GBDT	機械学習	0.488	6	17
複数系列 Random Forest	機械学習	0.506	7	-
Linear	機械学習	0.511	8	18
Ridge	機械学習	0.511	9	29
Random Forest	機械学習	0.513	10	12
LightGBM	機械学習	0.527	11	26
Auto ARIMA	統計	0.561	12	10
Comb4	コンビ	0.571	13	2
XGBoost	機械学習	0.604	14	19
Naïve 1	統計	0.612	15	-
SES	統計	0.612	16	11
GAM	機械学習	0.629	17	21
ARMA	統計	0.648	18	23
ARIMA	統計	0.651	19	14
SARIMA	統計	0.657	20	-
Lasso	機械学習	0.668	21	-
Elastic-Net	機械学習	0.668	22	27
DeepAR	機械学習	0.743	23	28
Theta	統計	0.761	24	5
SVM	機械学習	0.804	25	24
TBATS	統計	0.807	26	7
Decision Tree	機械学習	0.816	27	25
ETS	統計	0.912	28	9
Naïve 2	統計	0.940	29	15
Comb1	コンビ	0.980	30	6
Naïve S	統計	1.000	31	22
Prophet	統計	1.009	32	-
Holt	統計	1.294	33	16
RNN	機械学習	2.193	34	30
MLP	機械学習	3.383	35	31
Damped	統計	4.229	36	8

参考文献

- [1] Makridakis, S., Spiliotis, E., Assimakopoulos, V.: The M4 Competition: 100,000 time series and 61 forecasting methods, International Journal of Forecasting, 36, 54-74, 2020.
- [2] Sekitani, J., Murakami, H.: Framework for Comparing Accuracy of Time-Series Forecasting Methods, 2022 12th International Congress on Advanced Applied Informatics (IIAI-AAI), Kanazawa, Japan, pp. 669-672 (2022).