C-07

逐次学習を用いた電力需要予測モデルの性能評価

Performance Evaluation of Electricity Demand Forecasting Models Using Online Learning

田中 翔梧 †　　中本 幸一 †
Shogo Tanaka　　Yukikazu Nakamoto

1. はじめに

ある時間間隔に対する電力需要の予測モデルを構築する場合，予測対象の電力需要の傾向に合わない学習データが多くなると予測精度が劣化するため，定期的に予測モデルを再構築する必要がある．また，予測モデルを構築するごとに訓練データやハイパーパラメータなどを選択する作業が煩雑である．

こうした手間を削減するための 1 つの手段として，予測モデルを再構築せず 1 つ観測値が入力されるごとに更新を行う逐次学習（Online Learning）[7] という学習方法がある．しかし，従来の予測モデルの代わりとして，逐次学習で構築した予測モデルを用いるためには、予測モデルの性能を調査する必要がある．

そこで，本研究では逐次学習を用いて構築した電力需要予測モデルの性能を評価する．予測手法には Seasonal Autoregressive Integrated Moving Average with eXplanatory variables（SARIMAX）[3] を用いる．評価実験では，一階階差（1 時間前の観測値との階差）をとった系列における追加の説明変数の有無でそれぞれ学習と予測を行い，評価指標による予測モデルの評価を行う．

2. 手法

本研究は SARIMAX を用いて逐次学習を行うことで予測モデルを構築する（以下，逐次学習を行う SARIMAX を逐次 SARIMAX と呼ぶ）．ここで，SARIMAX は Seasonal Autoregressive Integrated Moving Average（SARIMA）[2] に説明変数を追加できる手法である．SARIMA には一般的な乗法（Multiplicative）SARIMA と非乗法（Nonmultiplicative）SARIMA の種類が存在するが [2]，本研究で用いる River[8] の SARIMAX は逐次 ARMA[1] をもとに実装されており，非乗法 SARIMA の実装になっていると思われる [11]．なお，River の SARIMAX のプログラムは本研究の実験方法に合わせて一部修正して用いる．さらに，本研究では逐次 SARIMAX の回帰手法として逐次線形回帰を用いる．したがって，本研究で用いる逐次 SARIMAX は式 (1) で表せる．

$$\Delta^d \Delta_s^D y_t = b + \varepsilon_t + \sum_{k=1}^{r} \beta_k x_{k_t} + \sum_{i=1}^{p} \phi_i \Delta^d \Delta_s^D y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{I=1}^{P} \Phi_I \Delta^d \Delta_s^D y_{t-sI} + \sum_{J=1}^{Q} \Theta_J \varepsilon_{t-sJ} \tag{1}$$

† 兵庫県立大学，University of Hyogo

ここで，Δ^d は d 階階差，Δ_s^D は周期 s の D 階季節階差，y は観測値，b は切片，ε は誤差項，r は追加の説明変数の次元，p は非季節における自己回帰（autoregressive）の次数（order），q は非季節における誤差の移動平均（moving average）の次数，P は季節における自己回帰の次数，Q は季節における誤差の移動平均の次数，β，ϕ，Φ，θ，Θ はそれぞれ係数を表す．

なお，本研究では最初から全観測値があるため，逐次学習に必要な階差と標準化を全観測値に対して行ってから予測モデルを学習できるが，実際に予測モデルを構築する際には，将来の観測値は利用可能でない．そのため，本研究では 1 つ 1 つのデータで学習する前に，階差と標準化といった学習のために必要な処理を逐次的に行う．Fekri ら [4] はデータに対して逐次的に正規化を行っている．

さらに，本研究では学習に必要な過去の観測値が利用可能な地点から逐次学習を開始する．しかし，学習開始地点によらず，学習開始時に過去に予測した誤差が利用可能でないという問題が生じるため，本研究では式 (1) で用いられている誤差の移動平均を用いないことにする．Nguyen ら [9] は乗法 SARIMA を用いた逐次学習を行っているが，本研究と同様に誤差の移動平均は用いていない．誤差の移動平均を用いないことで，式 (1) は簡潔に式 (2) で表せる．

$$\Delta^d \Delta_s^D y_t = b + \varepsilon_t + \sum_{k=1}^{r} \beta_k x_{k_t} + \sum_{i=1}^{p} \phi_i \Delta^d \Delta_s^D y_{t-i} + \sum_{I=1}^{P} \Phi_I \Delta^d \Delta_s^D y_{t-sI} \tag{2}$$

また，d 階階差と周期 s の D 階季節階差をとる系列を式 (3)，式 (4) にそれぞれ示す．

$$\Delta^d y_t = (1-B)^d y_t \tag{3}$$

$$\Delta_s^D y_t = (1-B^s)^D y_t \tag{4}$$

ここで，B はバックシフト演算子と呼ばれ，式 (5) で示す．

$$B^i y_t = y_{t-i} \tag{5}$$

式 (3) と式 (5) を用いると，本研究で用いる一階階差をとる系列は $\Delta y_t = y_t - y_{t-1}$ のようになる．

本研究では評価指標として Root Mean Squared Error（RMSE）を用いる．RMSE を式 (6) に示す．

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2} \tag{6}$$

ここで，N はデータ数，$\hat{y}$ は予測値を表す．

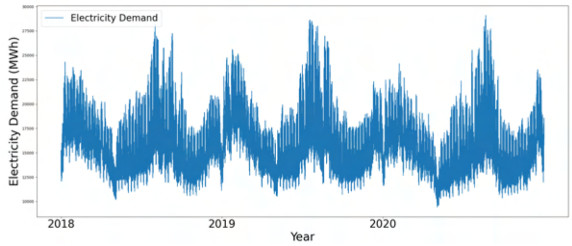

図 1　関西エリアの電力需要量（原系列）

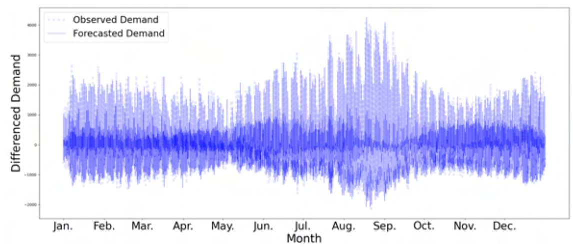

図 3　テスト期間の予測（休日変数なし）

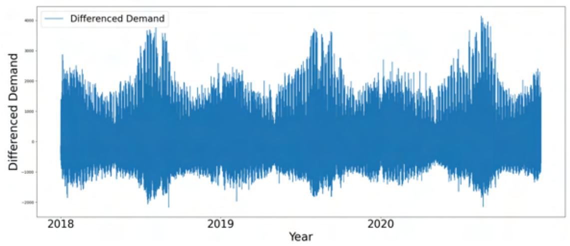

図 2　関西エリアの電力需要量（階差系列）

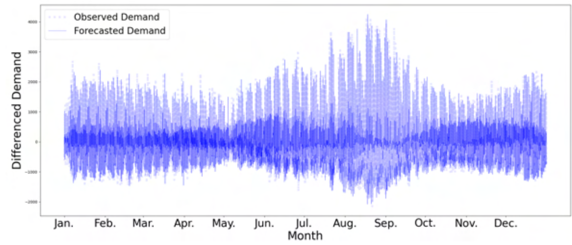

図 4　テスト期間の予測（休日変数あり）

3.　評価実験

3.1　実験設定

本研究で用いる電力需要データは関西電力送配電株式会社のホームページ [12] から，2018 年 1 月 1 日 0:00〜2020 年 12 月 31 日 23:00 までの関西エリアにおける電力需給実績の電力需要量を用いる．本電力需要データを図 1 に示す．本実験では，本電力需要データに対して一階階差をとった系列を用いる．本電力需要に対して一階階差をとった系列を図 2 に示す．

また，平日や休日の影響を考慮するために，追加する説明変数として休日変数を作成し用いる．休日変数を作成するにあたり，内閣府のホームページ [13] にある国民の祝日に関する情報をもとに，土日でない祝日と振替休日を土日に加えた日を休日とし，休日の時間を 1，それ以外の時間ならば 0 を値にもつ説明変数とする．なお，休日変数以外の説明変数は 1 つの観測値で学習する前に逐次的に標準化を行う．

本評価実験では，2018 年を準備期間，2019 年を検証期間（ハイパーパラメータ探索），2020 年をテスト期間とし，休日変数の有無でそれぞれの実験を行う．なお，休日変数を用いない条件では，逐次 SARIMAX ではなく逐次 SARIMA を用いたことになる．準備期間では逐次学習のみを毎時間の観測値を用いて逐次的に行う．なお，検証期間のハイパーパラメータ探索で最も過去の観測値が必要になる条件（具体的には一階階差をとり，かつ 6 週間前の観測値を学習で用いる場合）に合わせて，学習は 1,009 時間前の観測値を用いることができる 2018 年 2 月 12 日 1:00 から開始する．検証期間では準備期間で学習した予測モデルを用いて，引き続き学習は毎時間の観測値で逐次的に行い，予測を開始する．予測は毎日 0:00 の観測値で学習する前に 24 時間先（例えば 2019 年 1 月 1 日 0:00〜23:00）までの電力需要を予測する．そして，検証期間の観測値と予測値を用いて RMSE をもとに評価し，最適なハイパーパラメータの組み合わせを決定する．テスト期間では，検証期間で決定したハイパーパラメータの組み合わせを採用した予測モデルを用いて，引き続き検証期間と同様の方法で学習，予測，そして評価を行う．

3.2　ハイパーパラメータ探索

逐次 SARIMAX の設定するハイパーパラメータは各次数（p，P），周期（s），Stochastic Gradient Descent（SGD）[10] の学習率（lr）である．探索するハイパーパラメータの取りうる値の範囲を（p： 0〜6，P：0〜6，s：24，168，lr：0.01，0.001，0.0001）として，一階階差をとった系列に対して RMSE を用いて全探索を行う．この全探索は休日変数の有無によって別々に行う．また，階差処理は標準化と同様，準備期間から 1 つの観測値で学習する前に逐次的に行う．

ハイパーパラメータの全探索を行った結果，最も小さかった RMSE による評価値は，休日変数なしの条件で 216，休日変数ありの条件で 217 であり，選択されたハイパーパラメータの組み合わせはどちらも（p：0，P：6，s：168，lr：0.01）であった．ここで，この選択されたハイパーパラメータの組み合わせの説明として，6 つの観測値（1〜6 週間前）を用いて 0.01 の学習率で式 (2) を用いた逐次学習を行うことになる．休日変数の有無で，このハイパーパラメータの組み合わせを採用した予測モデルをテスト期間でそれぞれ評価する．

4.　実験結果と考察

テスト期間における評価実験の結果，RMSE の評価値は休日変数なしの条件で 222，休日変数ありの条件で 223 であり，どちらの条件も RMSE による評価値は検証期間からテスト期間を通して大きな変化はなかった．両方の条件での予測結果を図 3，図 4 にそれぞれ示す．図 3，図 4 はそれぞれ縦軸は階差をとった電力需要とその予測値，横軸は 2020 年の月を表し，薄い青の点線が階差をとった観測値，濃い青の点線が予測値を表す．RMSE による評価値や図 3，図 4 より，逐次 SARIMA や逐次 SARIMAX はモデル構築から数年であれば電力需要に対

して一定の誤差幅で安定した予測精度を保つことができることがわかった．

さらに，休日変数の有無で予測精度に差がなかったことから，休日変数は予測精度向上に対する効果がなかったことがわかる．これは，自己回帰の説明変数や階差をとったことによる影響力が大きく，休日変数のもつ影響力がほとんど打ち消されたためだと考えられる．このことから，逐次 SARIMAX で説明変数を追加して用いる場合，モデル構築前から自己回帰の説明変数や階差など多くの影響を考慮し追加する説明変数を決める必要がある．特に考慮したい大きな外部の要因がない限り，逐次 SARIMA を用いてハイパーパラメータの組み合わせを考慮するだけでも十分な可能性がある．

5. 関連研究

Nguyen ら [9] は Online Newton Step（ONS）[5] を用いた Online SARIMA（SARIMA-ONS）を提案した．時間単位の電力負荷（電力需要）データセットに対して，各時間に対して 1 つの予測モデルを構築し，24 個の予測モデルを用いて SARIMA-ONS の評価実験を行った．実験の結果，Mean Absolute Percentage Error（MAPE）において 24 個の予測モデルの平均で 5% 以下の予測精度を達成した．

Fekri ら [4] は深層学習と逐次学習を用いて電力負荷予測を行うために Online Adaptive Recurrent Neural Network を提案した．スマートメーターにより計測された時間単位の電力負荷データセットを用いて，温度や風速などの気象情報や年月日などの 11 個の説明変数で電力負荷を予測する実験を行った．実験の結果，Mean Squared Error（MSE）と Mean Absolute Error（MAE）において従来の Long Short Term Memory（LSTM）[6] やいくつかの標準的な逐次学習モデルに対して全体的に優れた予測精度を達成した．

これらの研究は予測精度を向上させるために，多くの予測モデルや説明変数を用いているが，実際にこれらの手法を用いて電力需要を予測するとなるとどちらも手間がかかる．

6. まとめ

本研究では，逐次学習を用いた電力需要予測モデルの性能を評価した．予測モデルとして SARIMAX を用いて，一階階差をとった系列における追加の説明変数の有無でそれぞれの条件における評価実験を行った．実験の結果，逐次 SARIMA と逐次 SARIMAX を用いると，本研究で用いた電力需要データで予測モデル構築時から数年であればある一定の誤差幅で安定して予測精度を保つことができることがわかった．しかし，電力需要に大きな影響を与えるようなイベントや災害などが生じた場合にどの程度逐次 SARIMA や逐次 SARIMAX が予測精度を保つことができるかわからないため，さらなる性能調査が必要だと考える．

さらに，説明変数を追加で用いる場合，自己回帰の説明変数や階差などの影響を分析して，慎重に説明変数を追加する必要があることがわかった．

より長期間または大きく変動がある電力需要データを用いたり，別の説明変数を追加で用いたりした実験により，逐次 SARIMAX を用いた予測モデルのさらなる評価や，別の逐次学習手法を用いた予測モデルの実験や電力需要データの特性を分析し実験に活かすことが今後の課題である．

参考文献

[1] Anava, O., Hazan, E., Mannor, S. and Shamir, O. : Online learning for time series prediction, Proc. the 26th Annual Conference on Learning Theory, pp.172-184 (2013).

[2] Box, G.E.P, Jenkins, G.M., Reinsel, G.C. and Ljung, G.M, : Time Series Analysis: Forecasting and Control, 5th ed., Wiley & Sons (2015).

[3] Cools, M., Moons, E. and Wets, G. : Investigating the Variability in Daily Traffic Counts through use of ARIMAX and SARIMAX Models: Assessing the Effect of Holidays on Two Site Locations, Transportation Research Record: Journal of the Transportation Research Board, Vol.2136, No.1, pp.57-66 (2009).

[4] Fekri, M.N., Patel, H., Grolinger, K. and Sharma, V.: Deep learning for load forecasting with smart meter data: Online Adaptive Recurrent Neural Network, Applied Energy, Vol.282, Part A (2021).

[5] Hazan, E., Kalai, A., Kale, S.: Logarithmic regret algorithms for online convex optimization. Machine Learning, Vol.69, pp.169–192 (2007).

[6] Hochreiter, S. and Schmidhuber, J.: Long short-term memory, Neural Computation, Vol.9, No.8, pp.1735-1780 (1997).

[7] Hoi, S.C.H., Sahoo, D., Lu, J., and Zhao, P.: Online learning: A comprehensive survey, Neurocomputing, Vol.459, pp.249-289 (2021).

[8] Montiel, J., Halford, M., Mastelini, S.M., Bolmier, G., Sourty, R., Vaysse, R., Zouitine, A., Gomes, H.M., Read, J., Abdessalem, T. and Bifet, A. : River: machine learning for streaming data in Python, Journal of Machine Learning Research, Vol.22, No.110, pp.1-8 (2021).

[9] Nguyen, Q.D., Nguyen, N.A., Tran, N.T., Solanki, V.K., Crespo, R.G. and Nguyen, T.N.A.: Online SARIMA applied for short-term electricity load forecasting, Research Square, Preprint (2021), DOI: https://doi.org/10.21203/rs.3.rs-439120/v1, (accessed 2022-07-14).

[10] Zinkevich, M.: Online Convex Programming and Generalized Infinitesimal Gradient Ascent, Proc. the Twentieth International Conference on Machine learning, pp.928-936 (2003).

[11] River: SNARIMAX (online), available from <https://riverml.xyz/dev/api/time-series/SNARIMAX/> (accessed 2022-07-18).

[12] 関西電力送配電株式会社 : 関西エリアの需給実績の公表 (オンライン), 入手先 <https://www.kansai-td.co.jp/denkiyoho/area-performance.html> (参照 2022-07-10).

[13] 内閣府 : 国民の祝日について (オンライン), 入手先 <https://www8.cao.go.jp/chosei/shukujitsu/gaiyou.html> (参照 2022-07-17).