

電力変圧器の油温予測

Electric Transformer Temperature ETT

寺本 雄飛

プロジェクト概要

電力変圧器の運用改善に関するPoCの目的は「油温予測モデルの作成」

目的

電力変圧器の油温を予測するモデルの作成
予測モデルが実運用に与える価値の提示

目指す姿

データを元に状態を監視し運用する
Condition Based Management(CBM)の実現

現状

経験則に基づいて単純な閾値監視
を中心に運用

課題内容

ETTデータセットのh1, h2を用いて、
2017-02-01 14:00:00 以降の油温予測を行う

行ったこと

EDA(探索的 データ解析)

データ特性の観察
季節性・トレンドの確認

予測のベースモ デル選定

EDA結果から適切な時系
列予測モデルを選定

実験計画の策定

モデルに基づき、評価指
標・評価方法を設計

実装・実験

実際にプログラムとして
実装・結果を確認

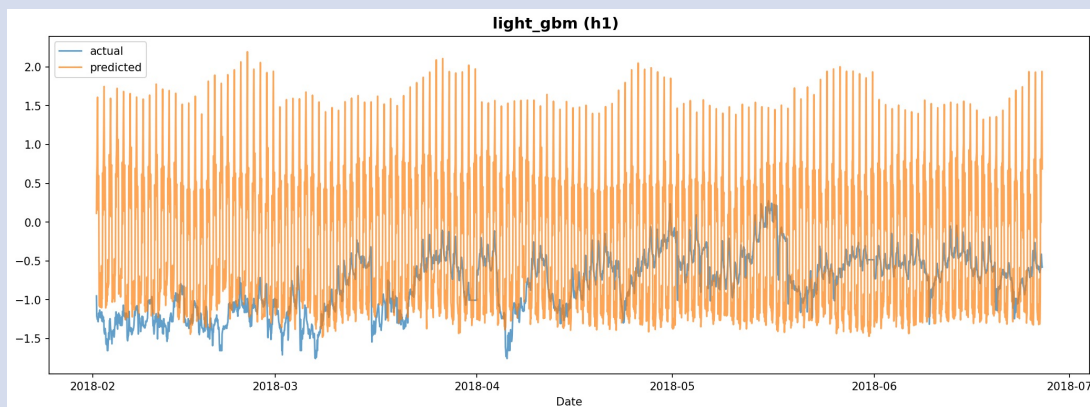
現状の整理

現状でできたこと・でき
なかったことを整理

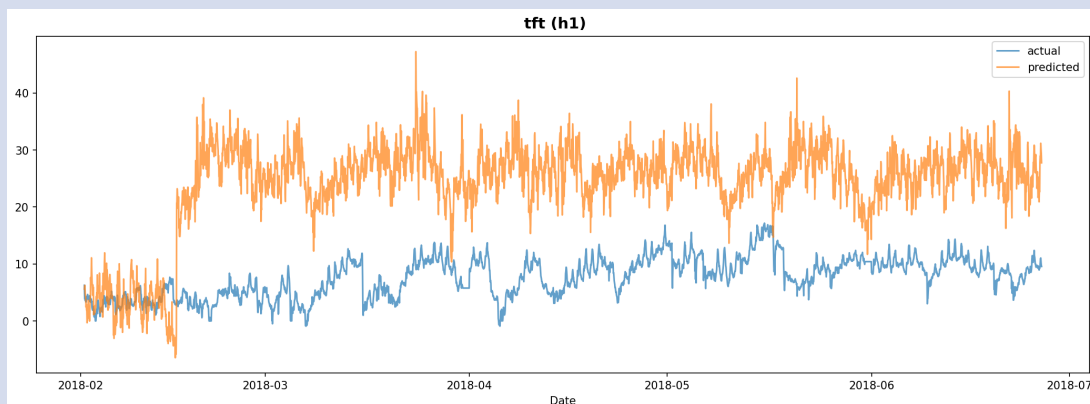
検証結果

- 油温をLightGBMとTemporally Fusion Transformer : TFTを用いて予測した
- モデルでRMSEがを超え、周期性の予測は可能だったものの、完全な予測には課題が残る

各種評価指標の比較



Metric	Value
MAPE	208.7
RMSE	1.067
MAE	0.8408
R^2	-6.575



Metric	Value
MAPE	308.8
RMSE	18.03
MAE	16.72
R^2	-26.39

EDAの方針とその結果について

データ概要

Electric Trasnformer Temperature : ETT(オープンデータ)
今回は1時間ごとのデータを予測対象として選択

- 2つの変圧器の油温と負荷状態(6種)のデータ
ETT-small-h1, ETT-small-h2
- 2016/07~2018/06の24ヶ月
- 欠損値なし

デモグラフィック情報

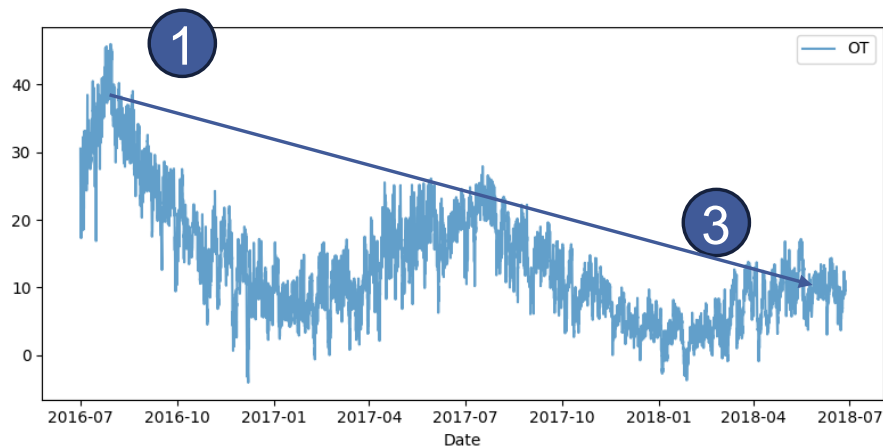
date	時刻情報 [datetime] 15分ごと(ETTM), 1時間ごと(ETTh)
OT	予測対象となる絶縁油温 [°C]
*UFL	有効負荷 (H: High, M: Medium, L: Low)
*ULL	無効負荷 (H: High, M: Medium, L: Low)

分析結果

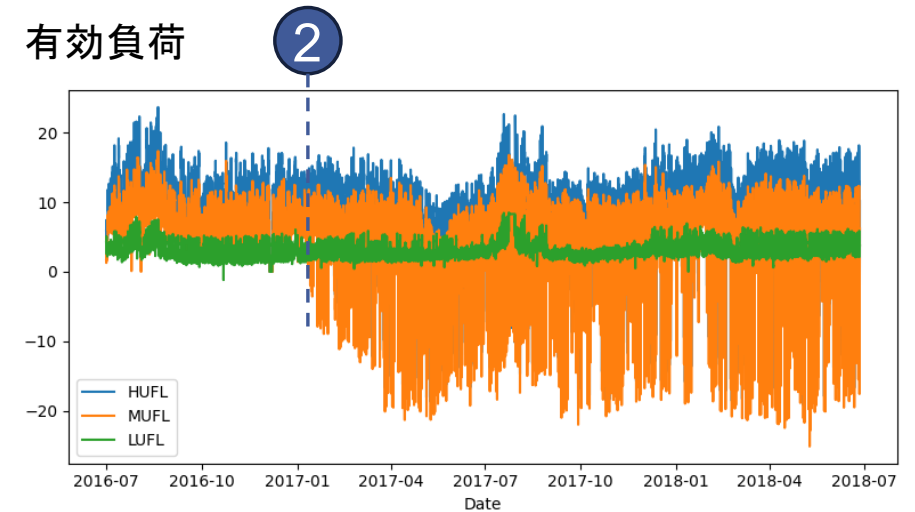
- | | | |
|--|---|--------------------------------------|
| ① OTは24時間・周期性が強く、*UFL, *ULLは強くない
年・月・週・日それぞれで周期性を持つ | ▶ | 周期性を分離した前処理を施し、予測モデルを構築する必要性が高い |
| ② ETT-h1に油温・電力負荷の異常が見られる
2017/01後半を境目に電力負荷特性が大きく変化 | ▶ | 異常発生前と後で電力負荷に対する油温の応答が変化している可能性がある |
| ③ ETT-h1の油温に年単位の低下トレンドが見られる
ETT-h2の油温には見られない | ▶ | 絶縁油の劣化により、電力負荷に対する油温の応答が変化している可能性がある |
| ④ ETT-h2に油温・電力負荷の異常が見られる
2017/04後半 | ▶ | 異常発生前と後で電力負荷に対する油温の応答が変化している可能性がある |

EDAの結果(ETTh1)

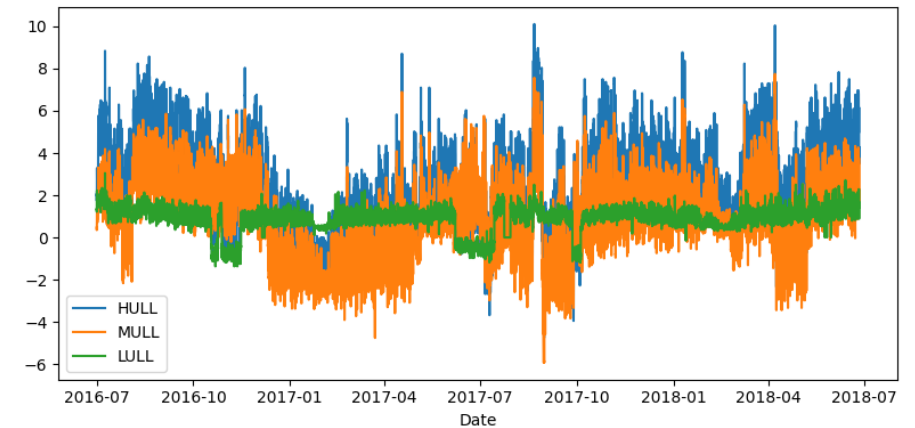
- OTは振動しながら減少するトレンドを持つ
- MUFLの下限值が2017/01から大きく低下
- 無効負荷は周期性が存在するが、周期が複雑に重なっていると見られる



有効負荷

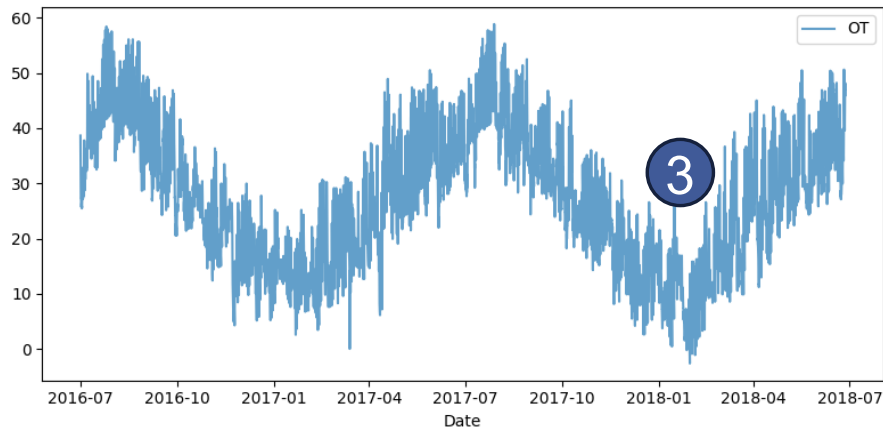


無効負荷

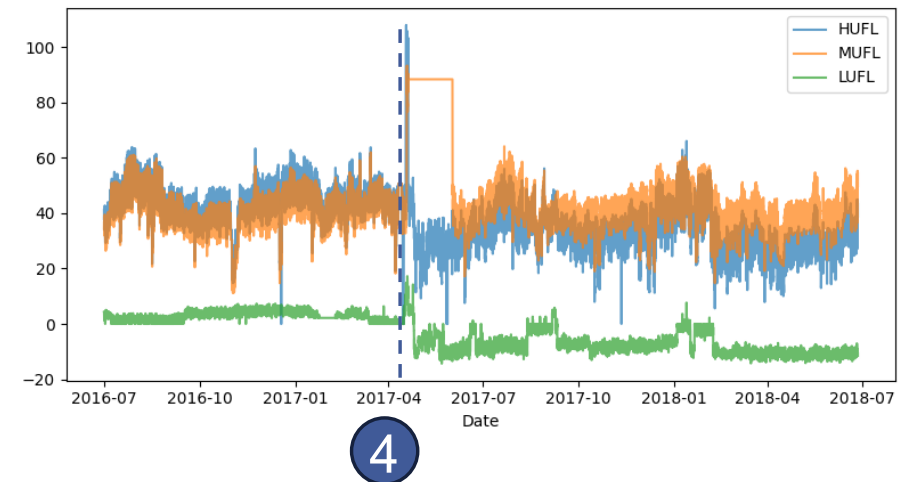


EDAの結果(ETTh2)

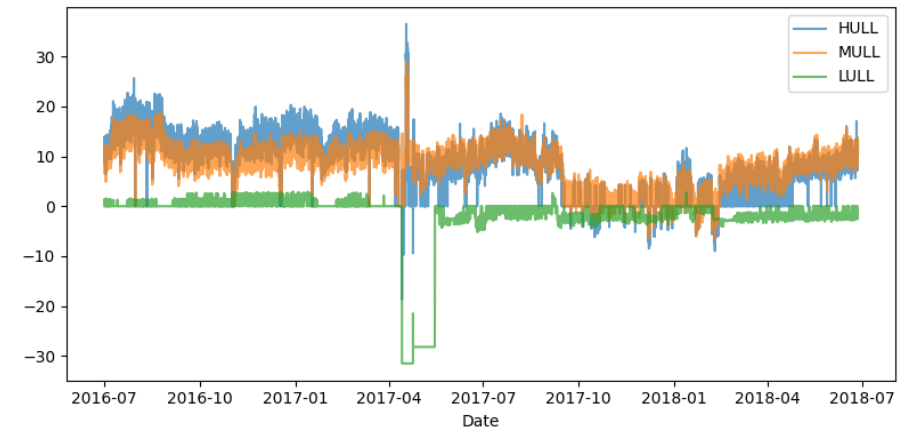
- OTは振動するものの、減少はしない
- MUFLの下限值が2017/01から大きく低下
- 無効負荷は周期性が存在するが、周期が複雑に重なっていると見られる



有効負荷

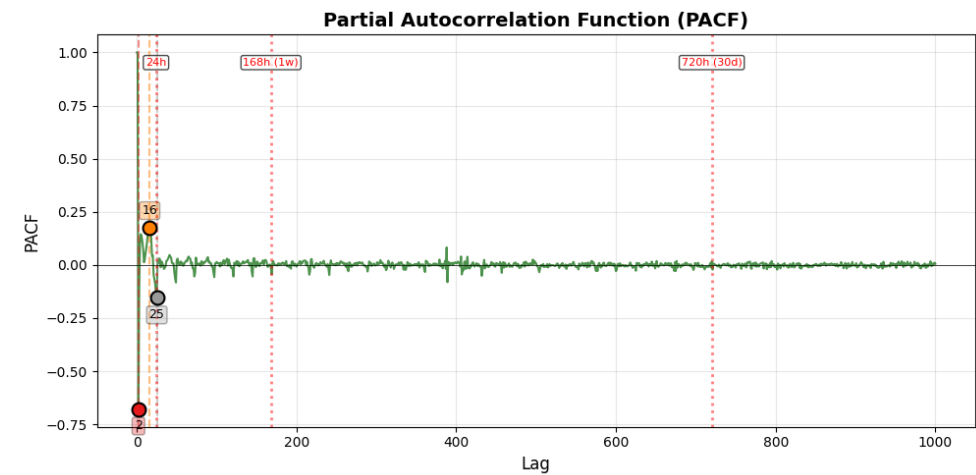
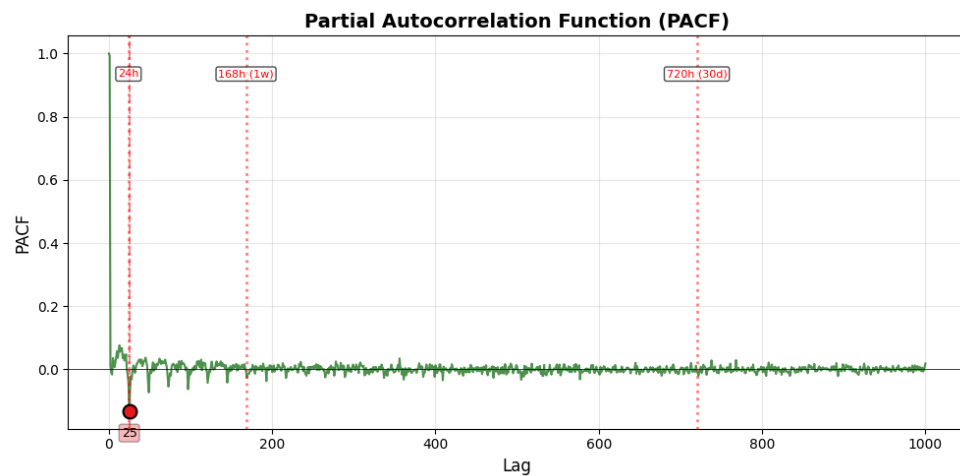
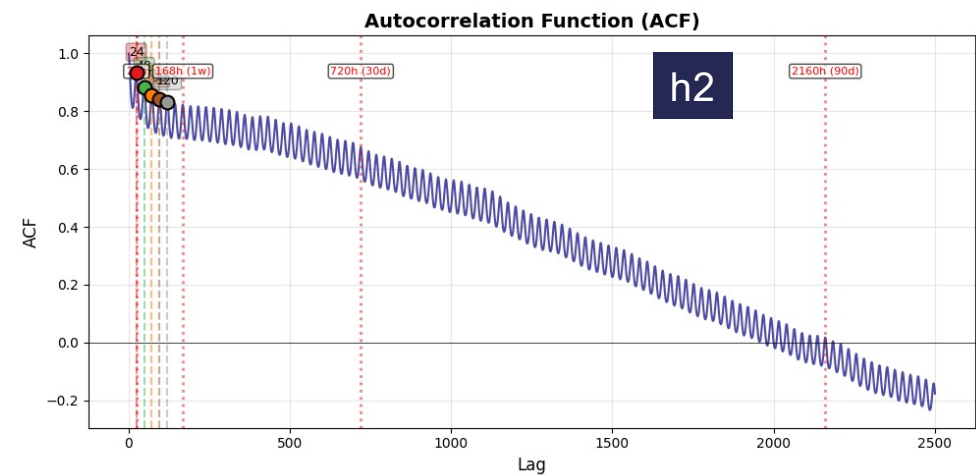
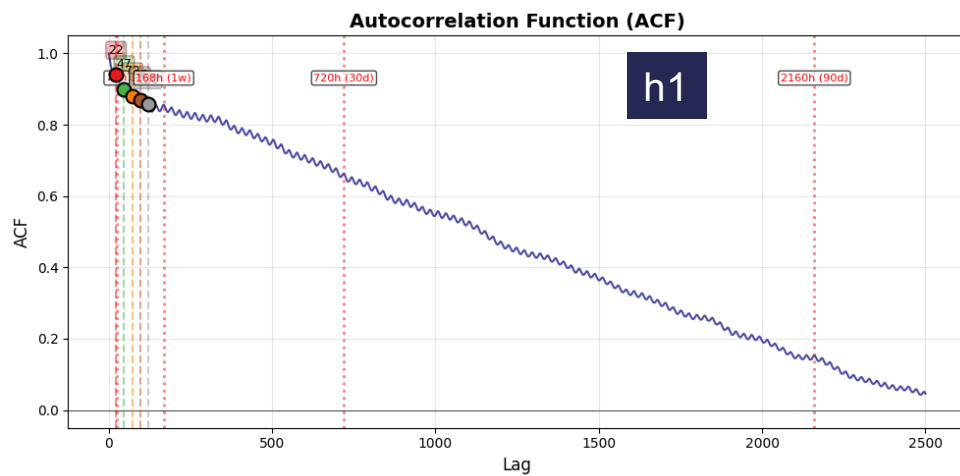


無効負荷



ACF, PACF(コレログラム)

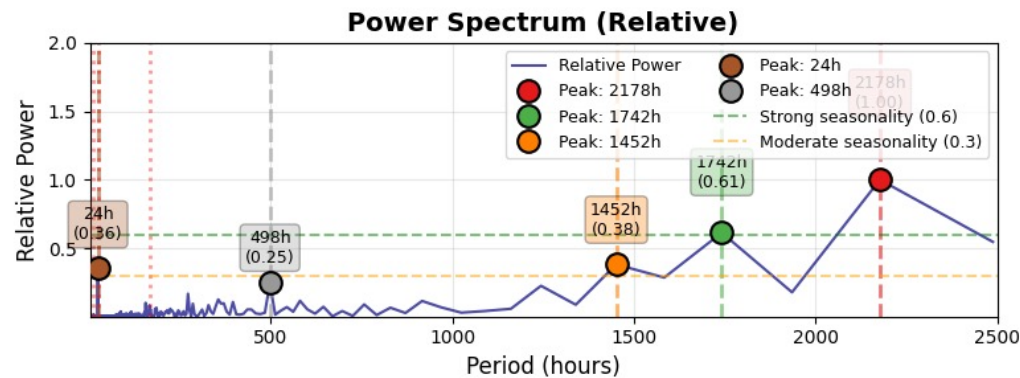
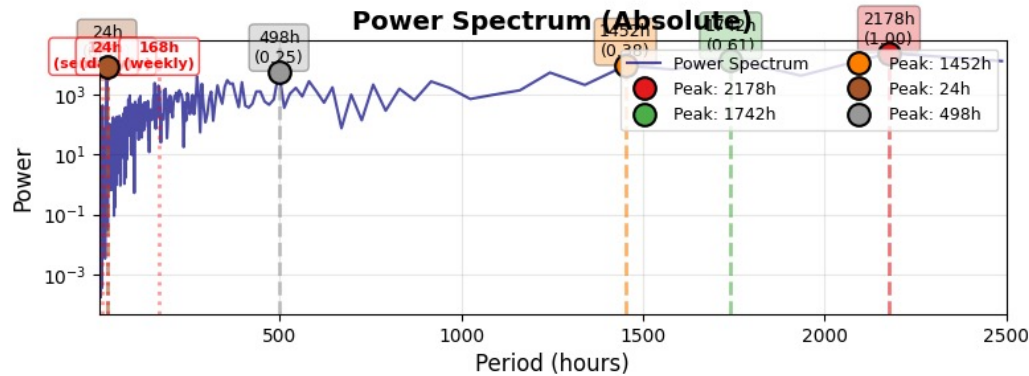
- ACFに24hの小さな振動を確認、48h、72hにも確認



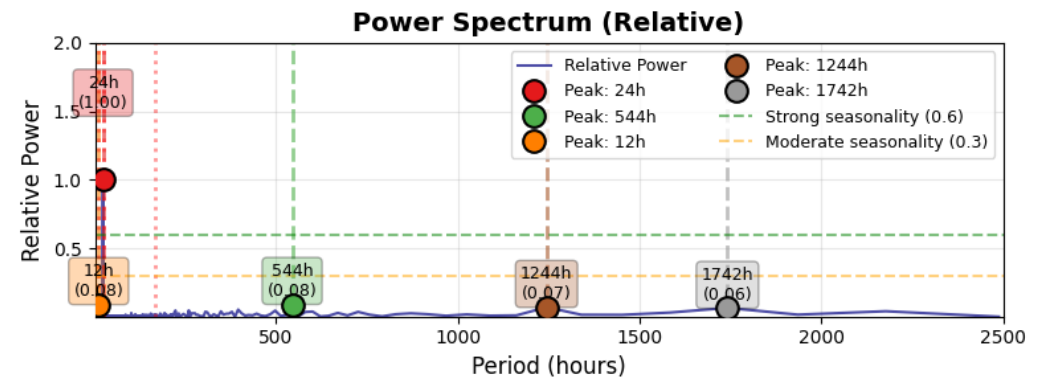
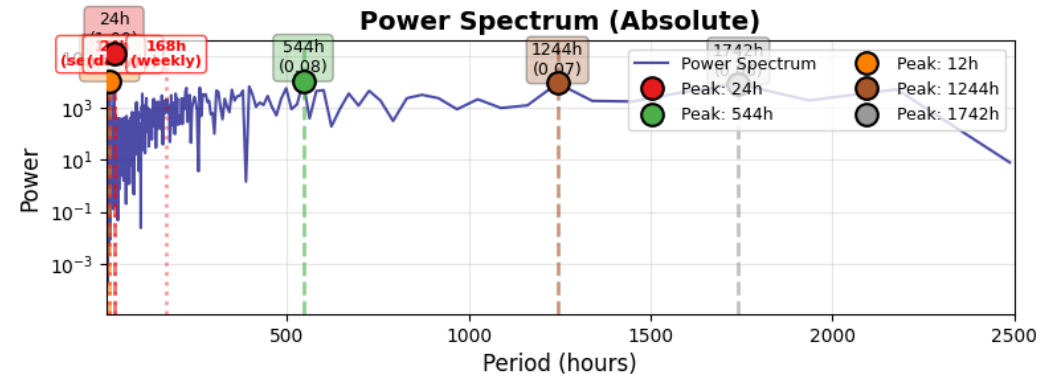
Power Spectrum

- パワースペクトル(フーリエ変換した時のある周波数成分の強さを表す)と周期の関係を調査
- OTは24h, 1742h, 2178hの成分が強いことが明確になり、分離周波数を特定
- *UFL, *ULLについても同様の検証を行い、周波数成分の特定を実施

ETTh1



ETTh2



データ前処理の手法

- EDAから周期性の分離の必要性が明らかになった
- Multiple Seasonal-Trend decomposition using LOESS MSTL分解を用いて周期性と変圧器の影響に分解

MSTL

EDAで観測された周期性を分離し、モデル精度を向上させる

$$R = Y - T - S_1 - S_2 - \dots - S_n$$

特徴

複数の周期性に分離する
季節性を分離することが可能

採用理由

周期性が混在しており、
それぞれの影響を分離することが必要なため

全ての特性に適用
OTのresidのみを予測対象とする

ラグ特徴量・移動平均

各種負荷に対してそれぞれ適用
OTに対しては使用せず

ラグ特徴量

t-n期前のデータを
特徴量として使用

移動平均

ある期間での平均を
特徴量として使用

特徴

観測前のデータに依存
する特性を反映できる

採用理由

EDAからデータの連続
性が強いことがわかって
いたため

sin cos特徴量

以下の式に沿った時間を基にした
特徴量

$$f_s = \sin\left(\frac{t}{period}\right), f_c = \cos\left(\frac{t}{period}\right)$$

採用理由

3時間程度の短い周期と季節・年単位の長い周期を
捉えるため

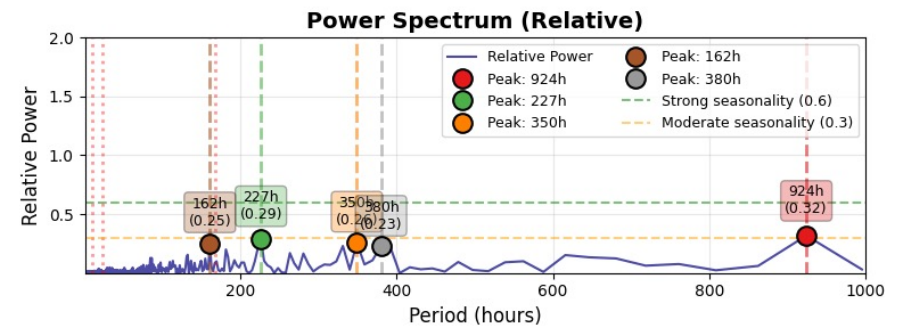
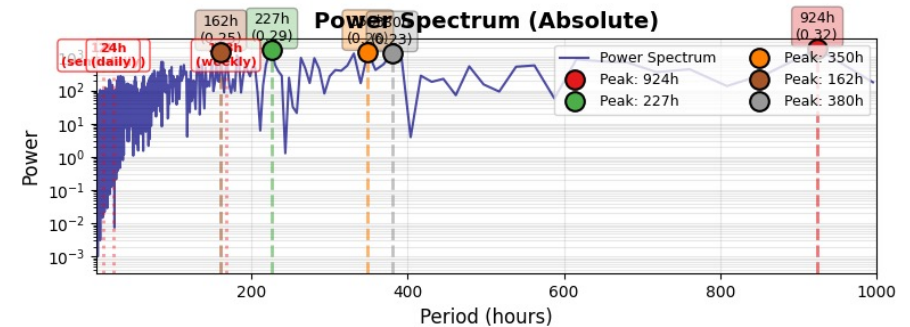
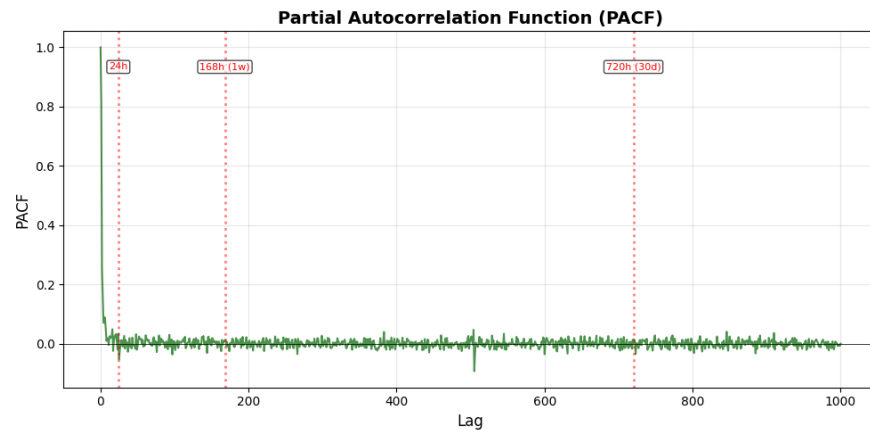
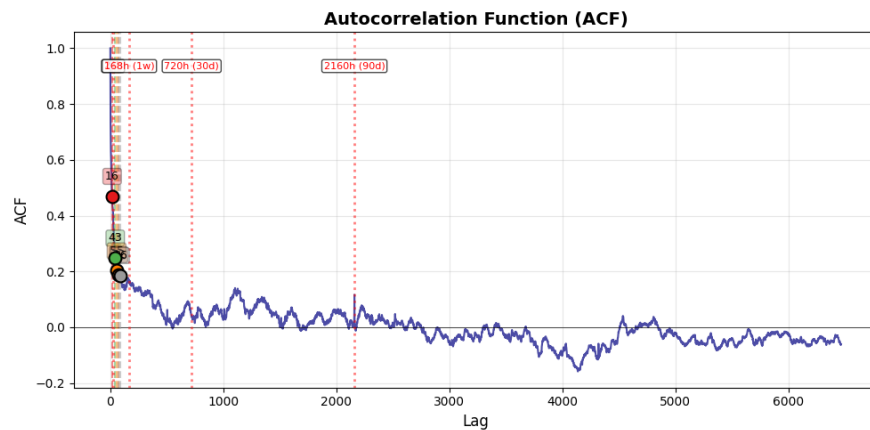
データの選別

データの除外

trainの先頭1000行分を除外
MSTLの最長周期の2160の半周期

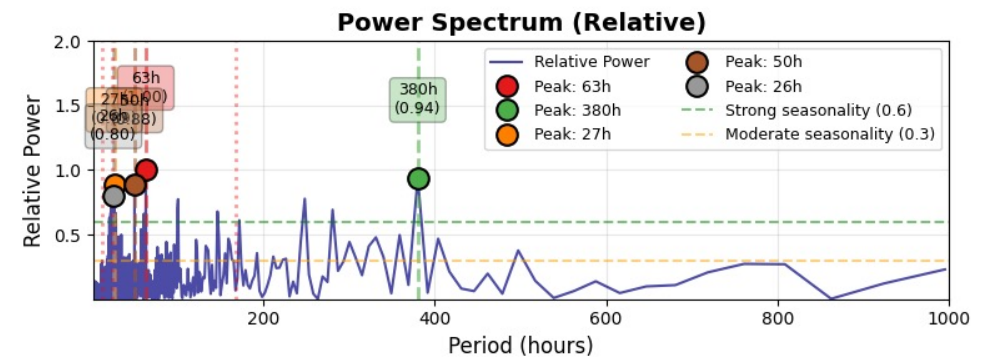
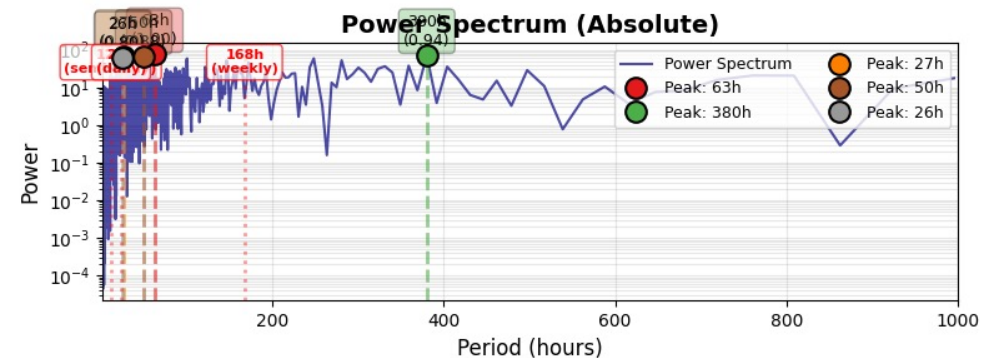
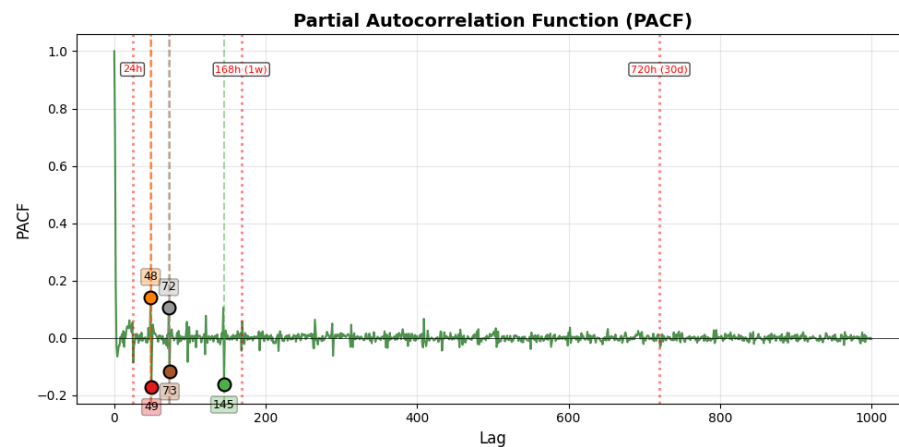
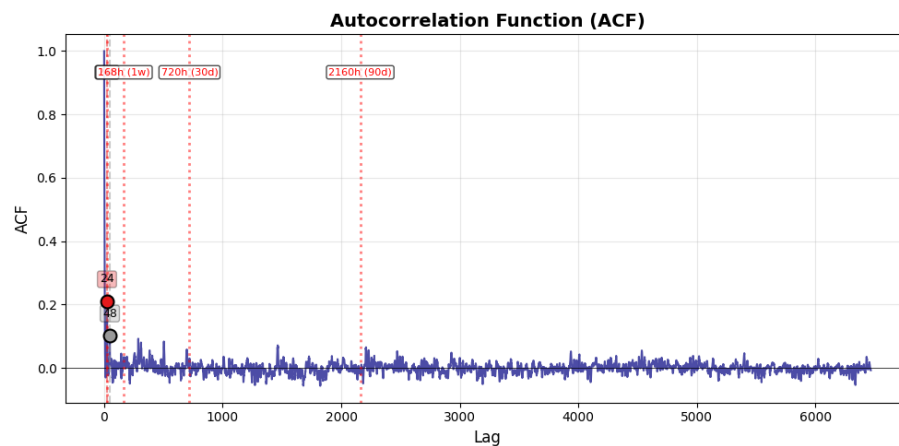
h1-OTの分離結果

- 前処理後の周波数特性について分析、全データはリポジトリの reports/figures/eda に掲載
- OTでは周期性が残存することが確認、予測の余地が見られる



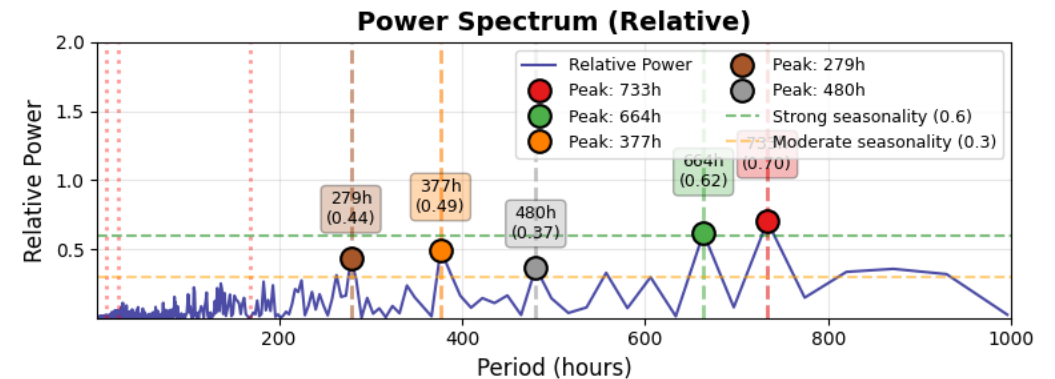
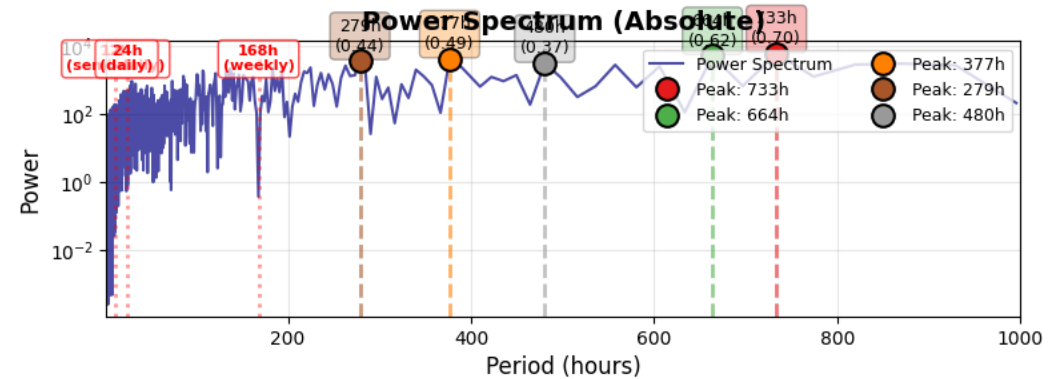
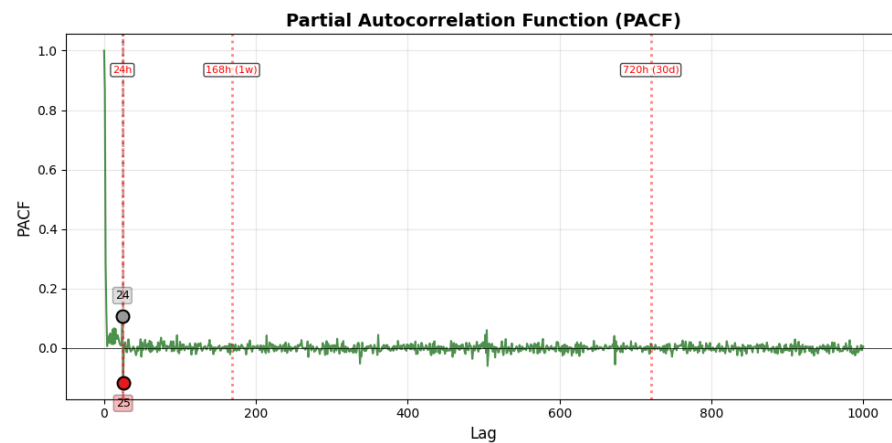
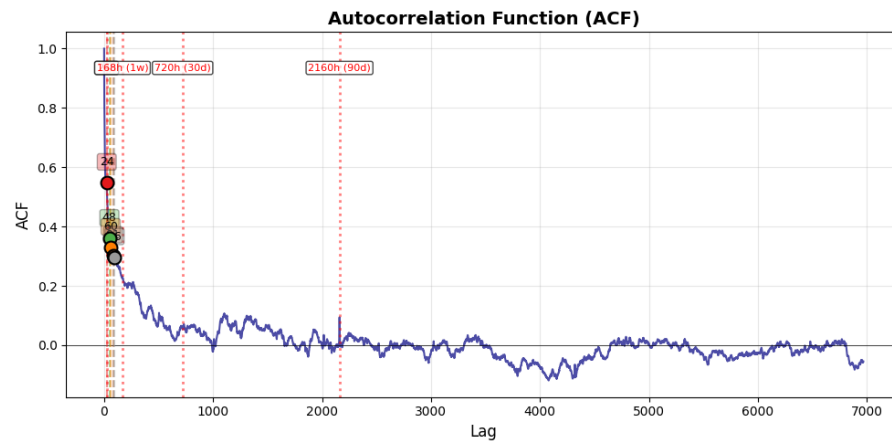
h1-HUFLの分離結果

- 前処理後の周波数特性について分析、全データはリポジトリの reports/figures/edaに掲載
- 他の変数では周期性の分離に成功、例としてHUFLを掲載



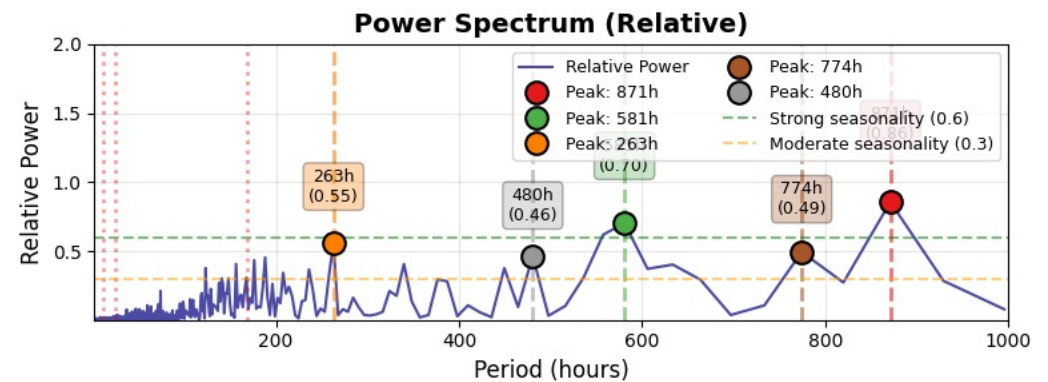
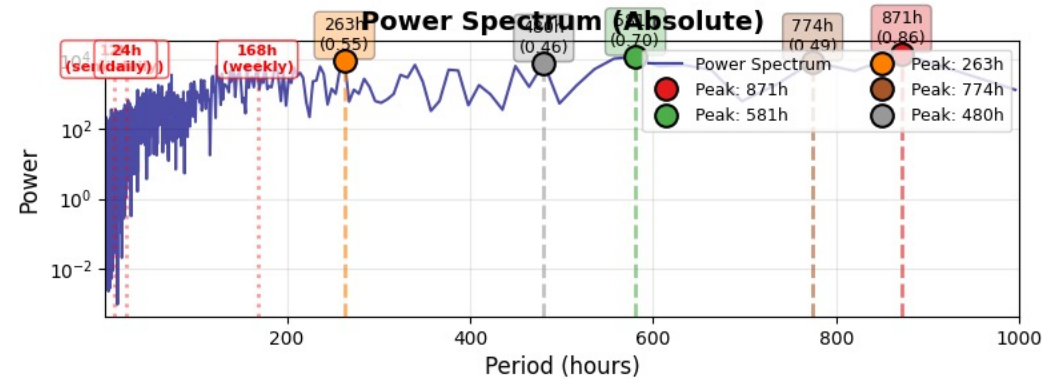
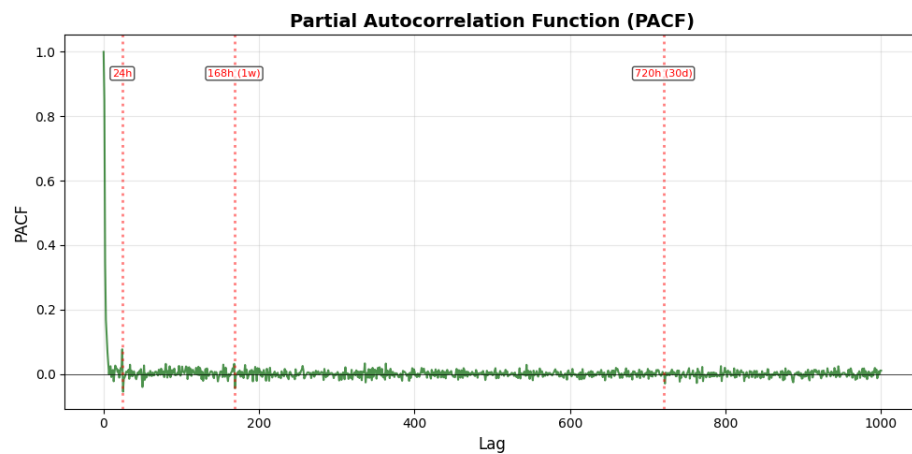
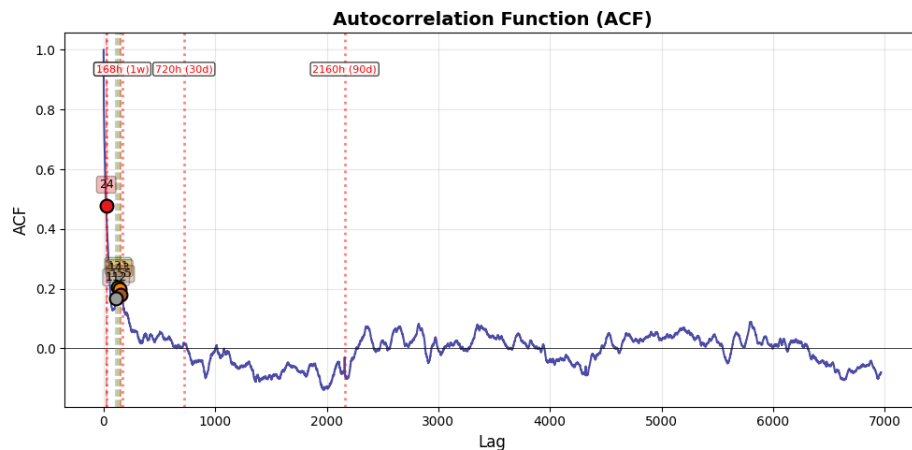
h2-OTの分離結果

- 前処理後の周波数特性について分析、全データはリポジトリの reports/figures/eda に掲載
- OTでは周期性が残存することが確認、予測の余地が見られる



h2-HUFLの分離結果

- 前処理後の周波数特性について分析、全データはリポジトリの reports/figures/eda に掲載
- HUFLでは周波数の分離に課題が残る



予測モデル設計

- Temporal Fusion Transformer : FTFとLightGBMの2モデルを用いて、予測を実行
- RMSEを用いてモデル学習を実行、MAPE・MAE・ R^2 を用いて予測精度を評価

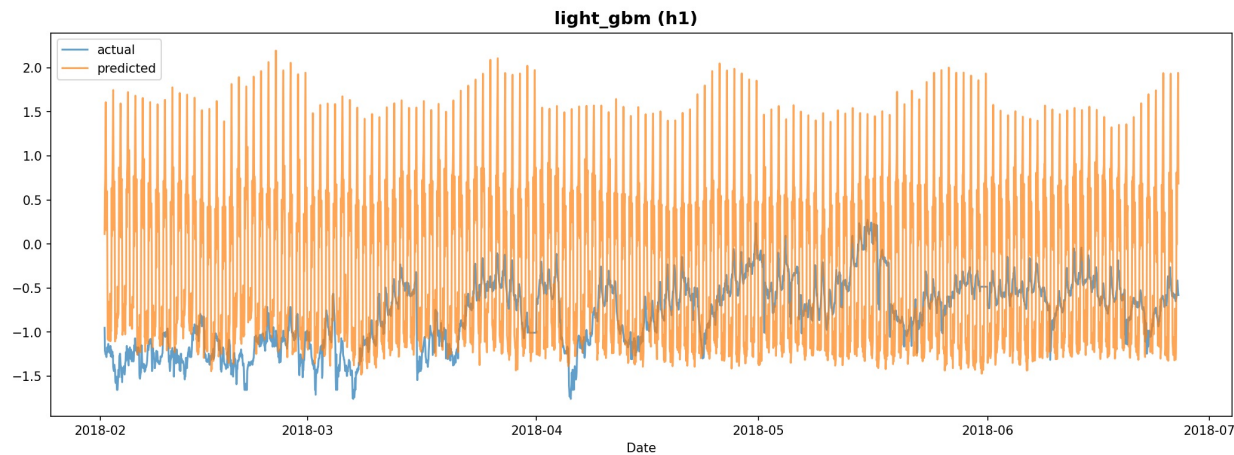
LightGBM	多変量回帰モデル 決定木ベースで高速
目的変数	resid, trend(OT)
説明変数	<ul style="list-style-type: none">*UFLを加工したMSTLデータsin cos特徴量
ハイパーパラメータの調整(Optuna)	
<ul style="list-style-type: none">num_leavesn_estimatorlearning ratefeature fractionbagging fractionbagging frequency	

TFT	Transformerベース 季節性を抽出可能
目的変数	resid, trend(OT)
説明変数	<ul style="list-style-type: none">*UFLを加工したMSTLデータ (seasonal, trend, resid)OTのseasonal, trend
ハイパーパラメータの調整(Optuna)	
<ul style="list-style-type: none">hidden sizeattraction headdropoutlearning rateLSTM Layer	

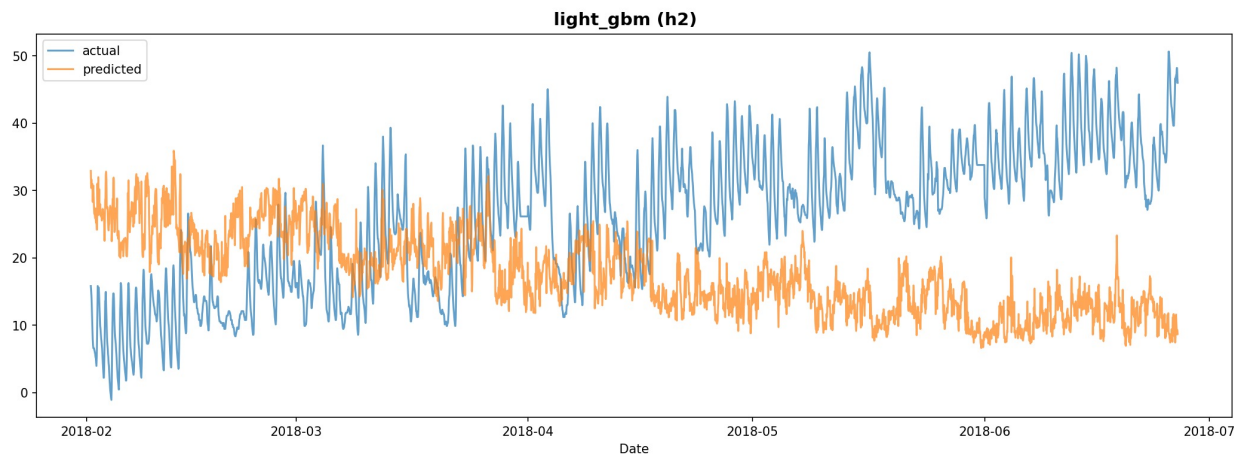
評価方法	
RMSE	Root Mean Squared Error 元データと同じ単位 モデルのvalidationに使用
MAPE	Mean Absolute Percentage Error 相対的な誤差の度合いを示す
MAE	Mean Absolute Error 元データと同じ単位の絶対誤差を示す
R^2	決定係数 予測精度の指標

LightGBMによる予測結果

- 周期に関する予測はある程度再現
- 絶対値が大きく異なる→MSTLのtrend・residualが再現できていないため



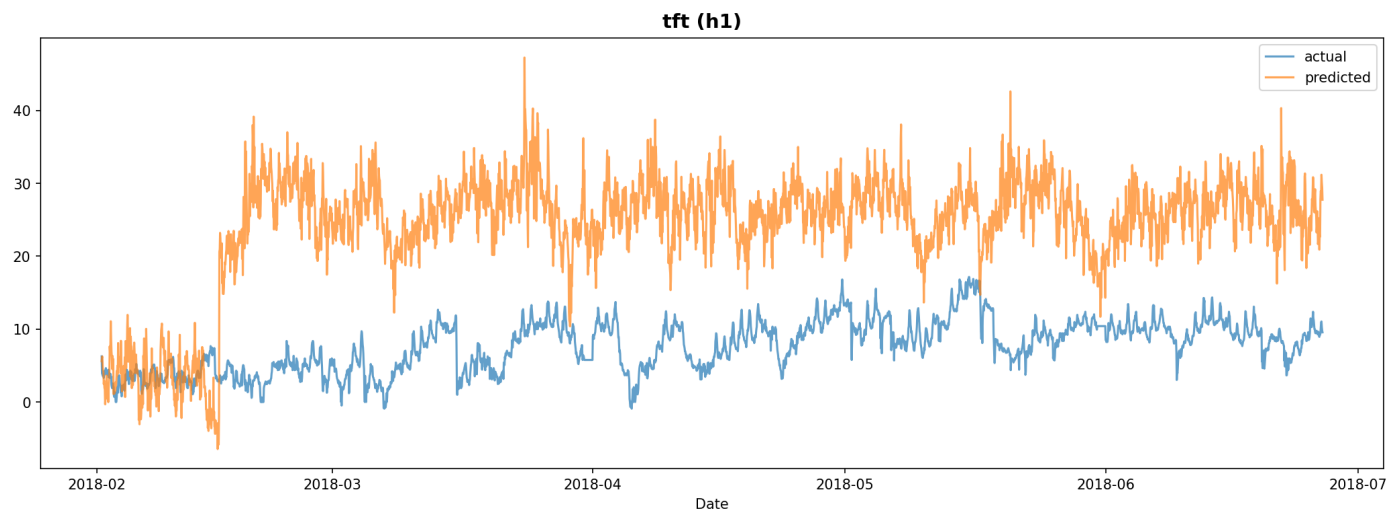
Metric	Value
MAPE	208.7
RMSE	1.067
MAE	0.8408
R^2	-6.575



Metric	Value
MAPE	164.1
RMSE	17.62
MAE	15.14
R^2	-1.714

TFTによる予測結果

- 周期に関する予測は概ね再現
- 計算時間の都合上、h1のみの結果
- 絶対値が大きく異なる→MSTLの全変数のtrend・residualが再現できていないため



Metric	Value
MAPE	308.8
RMSE	18.03
MAE	16.72
R^2	-26.39

工夫した点

1 EDAに時間をかけた

- 単純な時系列可視化
 - 周期性が存在することを確認
- ACF・PACF・パワースペクトルでの周期性の確認
 - どの周期が強く出ているのかを確認
 - 分離対象となる周期性を特定した

2 周期性の分離

- MSTLを用いて原系列を複数のseasonal項(S_n)とtrend項(T)、residual項(R)に分離
- trendとresidualの予測によってOTの予測を実現する方針を立てた

$$Y = R - T - S_1 - S_2 - \dots - S_n$$

3 複数モデルによる評価

3モデルを使用し、適切なモデルを選定

- SARIMAX(今回は精度が低く不採用)
 - 季節性の考慮
- LightGBM
 - 基礎的な予測モデル
- TFT
 - Transformerベースの時系列モデル

4 コードの保守性を重視

コードの再利用・実験の簡易化を実現

可読性

ドキュメントを充実させる
docstringを細かく記述

型安全性

pydanticを用いて型指定とvalidation
を実行

課題と今後の展望

課題

異常値と正常値の判定	ドメイン知識の不足により、異常と正常の判定が実現できていない →モデル向上の余地 h1 : 2017/01ごろのMUFLの異常減少 h2 : 2017/04ごろの
特徴量の不足	電力負荷の特性や外気温などの変圧器外の情報が不足しているため、精度が低下
トレンドの再現	異常値が混在しているためにMSTLのトレンドを近似できず、テストデータに外挿できない
真の予測とは言えない	*UFL, *ULLなどのデータを観測済みとしているため、完全な予測モデルとはいえない
実行時間	モデルが重く、計算時間が長い

今後の展望

トレンドの再現	<ul style="list-style-type: none">多段予測モデルを実装電力負荷のresid, trendを予測し、testに外挿→OTの予測平滑化により、再現性を向上<ul style="list-style-type: none">log, exp
異常値の抽出	<ul style="list-style-type: none">ドメイン知識の補完によって異常値と正常値を区別異常値から正常値を特定 →異常検知を用いてCBMの実現に貢献 DAEなどのNN系異常検知モデルの使用
アンサンブル	長期予測でトレンドを予測、間を短期予測で補完 <ul style="list-style-type: none">長期予測(NN系モデル) 月・年単位短期予測(線形モデル) 24h, 48h単位



Appendix

追加資料



特徴量セット

- 詳細はリポジトリのconfig/exp_004.yml, exp_005.ymlを確認

MSTL

periodsのみを指定
seasonalとtrendに関しては自動で設定

h1

h2

Feature	Periods	Feature	Periods
OT	24, 504, 720, 2160	OT	24, 336, 504, 2160
HUFL	12, 24 48, 72	HUFL	12, 24, 168, 2160
HULL	12, 24, 48, 504	HULL	12, 24, 504, 720
MUFL	12, 24, 504	MUFL	12, 24, 504, 720
MULL	12, 24	MUFL	12, 24, 504, 720
LUFL	12, 24, 48, 720	MULL	12, 24, 480
LULL	12, 24	LUFL	12, 24, 720
		LULL	12, 24

ラグ特徴量(Lag)・移動平均(MA)

MSTLの結果から分離できなかった周期を補完

h1

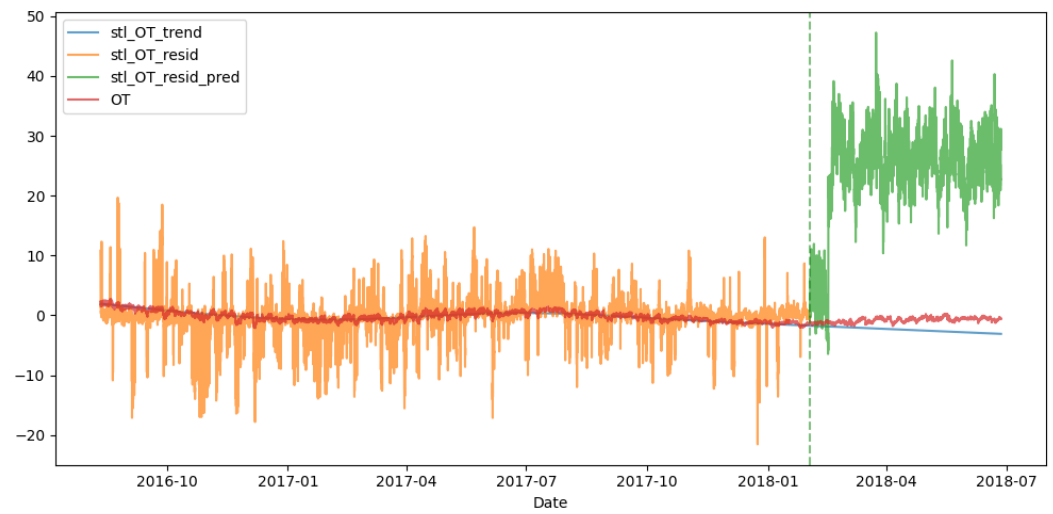
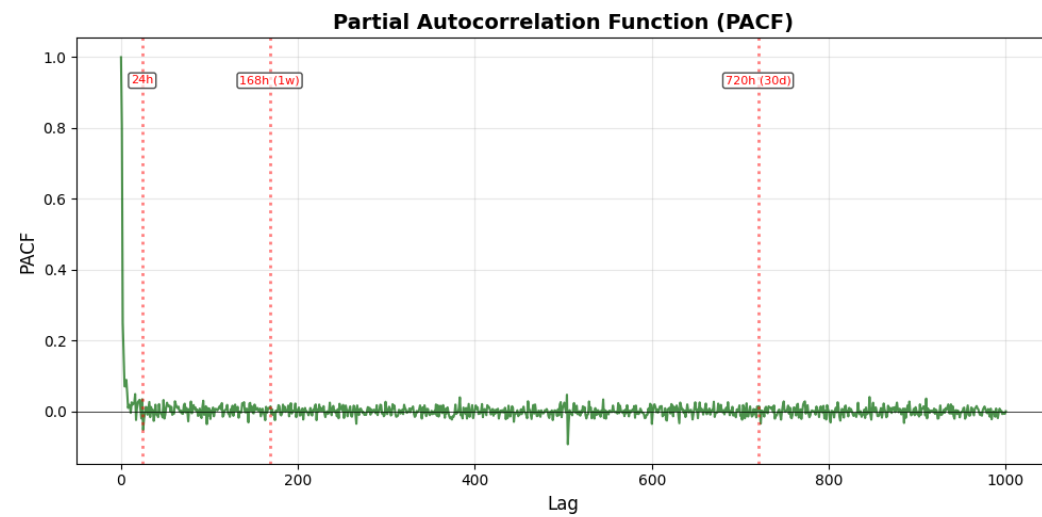
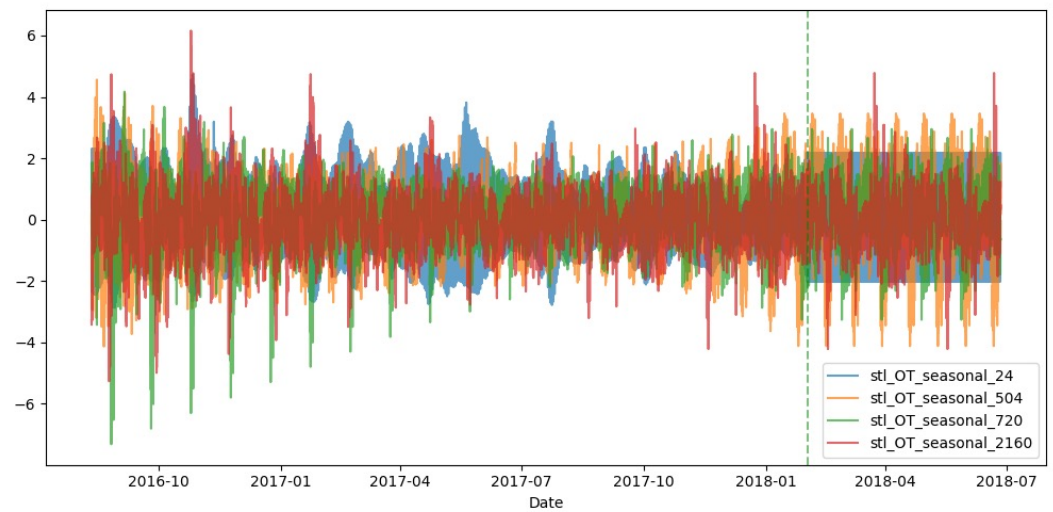
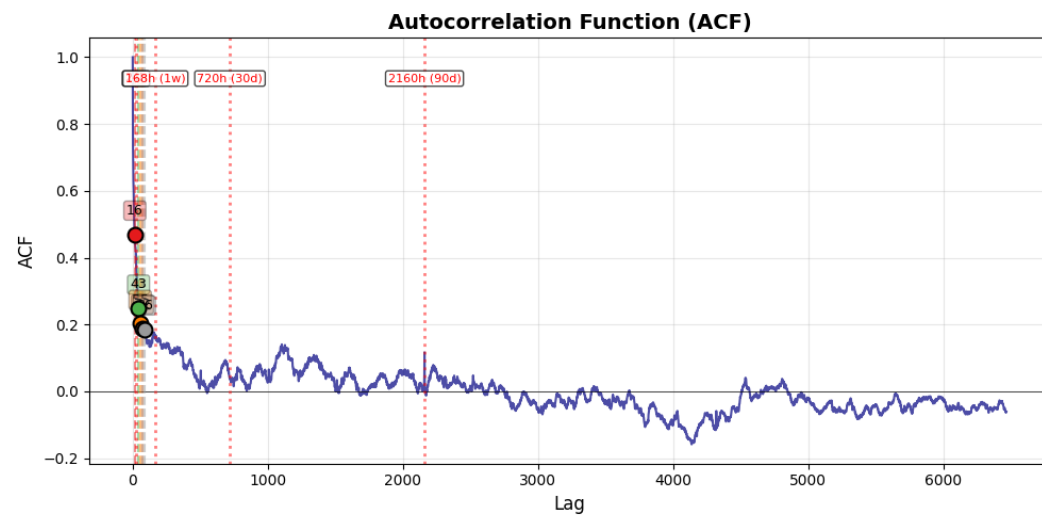
h2

Lag: 6, 12, 24, 168

Lag: 6, 12, 24, 168

Feature	MA windows	Feature	MA windows
OT	3, 12, 24, 168, 720	OT	3, 12, 168, 720, 4800
HUFL	3, 12, 24, 168	HUFL	3, 24, 48, 168
HULL	12, 24, 504, 720	HULL	3, 24, 240, 168, 504
MUFL	3, 72, 168, 720	MUFL	3, 12, 24, 168, 720
MULL	3, 72, 168	MULL	3, 12, 24, 168, 720
LUFL	3, 168, 720	LUFL	3, 12, 24, 168, 720
LULL	3, 72	LULL	3, 12, 24, 168, 720

h1-OTの分離結果



h1-HUFLの分離結果

